Adaptive State Allocation Algorithm in MLSD Receiver for Multipath Fading Channels: Structure and Strategy

Hossein Zamiri-Jafarian and Subbarayan Pasupathy, *Fellow, IEEE*

***Abstract*— This paper presents the adaptive state allocation (ASA) algorithm, a new scheme based on maximum likelihood sequence detection (MLSD) of signals transmitted over Rayleigh fading channels. Although MLSD is an optimal scheme, its computational complexity limits many applications. The ASA algorithm is a detection method whose performance is close to that of MLSD, but with greatly reduced computational complexity. Adaptive state partitioning in the trellis diagram is used in this algorithm by measuring the short-term received signal power. Also, an adaptive threshold for selecting only a few states of the trellis is employed in this algorithm based on the Chernoff distance between the probability density functions (pdf's) of correct and incorrect branch metrics. ASA-DF, a special case of ASA using decision feedback, shows very good tradeoff between performance and computational complexity for selective fading channels. Using ASA with diversity reception not only improves the performance, but also decreases the computational complexity in comparison with the computational complexity of using MLSD with diversity reception.**

***Index Terms*— Adaptive signal detection, complexity theory, fading channels, maximum likelihood detection, multipath channels, Rayleigh channels.**

I. Introduction

COMMUNICATION over fading multipath channels such as land-mobile radio and indoor wireless communication suffers from intersymbol interference (ISI) and fading phenomena. Both ISI and fading imply correlation among a sequence of symbols which are received by the receiver. It is well known that maximum likelihood sequence detection (MLSD) exploits these correlations and is the optimal receiver for these channels in the sense of minimizing the sequence error rate for equiprobable sequences [1]–[3].

Forney [4] proposed the MLSD structure based on the Viterbi algorithm as the optimal receiver for channels with ISI and additive white Gaussian noise. The MLSD receiver for Rayleigh flat fading channels and continuous phase modulation was presented by Lodge and Moher [1]. Dai and Shwedyk [2] developed sequential sequence estimation for selective Rayleigh fading channels. Using the innovations approach, an MLSD algorithm which demodulates the received signal recursively was proposed by Yu and Pasupathy [3].

The complexity of MLSD motivated a considerable amount of research to find other algorithms or structures whose performances are close to the optimum with less complexity. One group of efforts concentrated on using an equalizer to reduce the duration of the overall impulse response of the channel/equalizer such that the number of states in the trellis diagram decreases [5]–[9]. Qureshi and Newhall [5] used the linear equalizer to shape a desired overall impulse response, and in [6] and [7] the optimized desired impulse response was investigated. Lee and Hill [8] and Wesolowski [9] proposed to use decision feedback equalization (DFE) to shorten the overall impulse response.

Another group of efforts focused on reducing the computational complexity of the trellis by selecting only a subset of the total number of states. There are two main strategies for selecting the subset. The first strategy selects some states which are in the trajectories of the more likely correct paths [10]–[14]. The M algorithm [15] is perhaps the most famous among these algorithms and serves as the main core for other algorithms using this strategy. At every stage in the trellis diagram, the M algorithm allows only M paths with lower costs to extend into the next stage. Tradeoff between computational complexity and performance of the M algorithm is controlled by the value of M. Anderson [12] evaluated the least M in the M algorithm for decoding of convolutional codes. Simmons [13], [14] proposed to select the states whose costs were below a fixed value.

The second strategy proposed in [16]–[19] is known as reduced-state sequence estimation (RSSE). In this method, by using the Ungerboek set partition, only a fixed subset of states is selected and, in contrast to the M algorithm, the subset in RSSE is always fixed. The delayed decision-feedback sequence estimator in [18] and the decision-feedback sequence estimator (DFSE) in [16] are special cases of RSSE. In contrast to conventional DFE, where only one detected sequence of symbols is considered as a decision sequence for detecting the new received symbol, more than one detected sequence are considered in DFSE as decision sequences. The idea of optimum partitioning, based on keeping the minimum distance error rate of the partitioned set close to that of the unpartitioned set, is considered in [20] and [21].

In this paper, after deriving the MLSD structure for fading channels (both frequency nonselective and selective) without

Manuscript received June 18, 1996; revised September 4, 1997. This work was supported by a grant from the Canadian Institute for Telecommunications Research under the NCE program of the Government of Canada.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ont., M5S 3G4, Canada (e-mail: zamiri@comm.utoronto.ca).

Publisher Item Identifier S 0018-9545(99)00716-1.

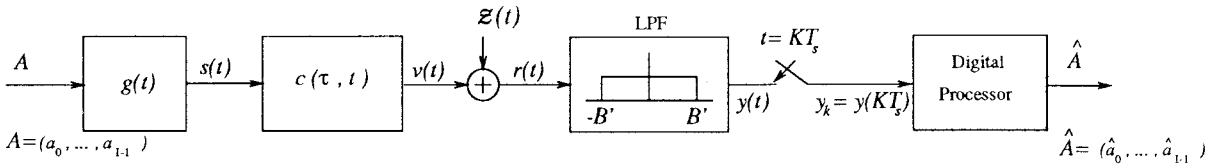


Fig. 1. The communication system over fading multipath channels.

restriction on the fading and signaling scheme, we propose a new algorithm called the adaptive state allocation (ASA) algorithm which greatly reduces the computational complexity of the MLSD receiver with a performance close to the optimum. With an adaptive state partitioning, based on measuring the short-term received signal power, the states of the trellis diagram are fused or diffused in the ASA algorithm. Also, some states are selected as the more likely correct states by using an adaptive threshold. We introduce ASA with decision feedback (ASA-DF) which has a good tradeoff between performance and computational complexity in multipath fading channels as well. ASA-DF can control the number of multiplications in each branch metric by using decision feedback.

The organization of this paper is as follows. After an introduction, the models of Rayleigh fading channels are described in Section II. In Section III, the MLSD algorithm is derived. The structure, strategy, and implementation of the ASA algorithm are considered in Section IV. Diversity reception is applied to the MLSD and the ASA algorithms in Section V. Computer simulation results and comparisons are presented in Section VI. Finally, Section VII presents some conclusions.

II. Channel Models for Band-Limited Signals

It is well known that fading multipath channels can be modeled as time-variant systems. The equivalent low-pass impulse response $c(\tau,t)$ represents the output of the fading multipath channel at time t due to an input impulse applied at time $t-\tau$. Based on the central limit theorem, $c(\tau,t)$ is a complex-valued Gaussian random process. When there is no fixed path between the transmitter and the receiver, the mean value of $c(\tau,t)$ is zero and $|c(\tau,t)|$ at any instant t is Rayleigh distributed and the channel is referred to as a Rayleigh fading multipath channel [22], [23].

A communication system over the fading multipath channel is shown in Fig. 1. The equivalent low-pass transmitted signal is

$$s(t)=\sum_{i=0}^{\lfloor t/T\rfloor} a_i g(t-iT) \tag{1}$$

where $\{a_i\}_{i=0}^{I}$ are independent identically distributed equiprobable symbols taken from a q-ary constellation, T is the symbol period, $g(t)$ is the impulse response of the transmitter filter, and $\lfloor\cdot\rfloor$ is the floor function (the largest integer not exceeding the argument). The received waveform is

$$r(t)=v(t)+\mathcal{Z}(t)=\int_{-\infty}^{+\infty} c(\tau,t)s(t-\tau)\,d\tau+\mathcal{Z}(t) \tag{2}$$

where $c(\tau,t)$ is the equivalent low-pass impulse response of the fading multipath channel and $\mathcal{Z}(t)$ is a circularly symmetric zero-mean white complex Gaussian random process with an autocorrelation function $R_{\mathcal{Z}}(\tau)=\hat{N}_0\delta(\tau)$ [24]. By using the sampling theorem when the bandwidth of $s(t)$ is B, $v(t)$, the signal component in $r(t)$ becomes

$$\begin{aligned} v(t)&=\int_{-\infty}^{+\infty} c(\tau,t)\sum_{l} s(t-lT_s)\,\mathrm{sinc}\left(\frac{\tau-lT_s}{T_s}\right)d\tau \\ &=\sum_{l} s(t-lT_s)\hat{h}_l(t) \end{aligned} \tag{3}$$

where $\mathrm{sinc}(x)=\frac{\sin(\pi x)}{\pi x}$, $T_s\le\frac{1}{2B}$ is the sampling period and

$$\hat{h}_l(t)\doteq\int_{-\infty}^{+\infty} c(\tau,t)\,\mathrm{sinc}\left(\frac{\tau-lT_s}{T_s}\right)d\tau. \tag{4}$$

If f_d is defined as the maximum Doppler frequency of the multipath fading channel, the power spectral bandwidth of $\hat{h}_l(t)$ is limited to f_d. Thus, from the statistical viewpoint, $h_l(t)$, the output of $\hat{h}_l(t)$ filtered by an ideal low-pass filter of bandwidth f_d, is equal to $\hat{h}_l(t)$

$$\begin{aligned} h_l(t)&\doteq\int_{-\infty}^{+\infty} 2f_d\hat{h}_l(\lambda)\mathrm{sinc}(2f_d(t-\lambda))\,d\lambda=\hat{h}_l(t), \\ &\text{for all } l. \end{aligned} \tag{5}$$

When the coherence bandwidth $(\Delta f)_c$ of the fading multipath channel is $(\Delta f)_c>2B$, the channel is called frequency-nonselective (or flat) fading (FNF), and when $(\Delta f)_c<2B$, the channel is called frequency-selective fading (FSF). Since the essentially nonzero duration of the autocorrelation function of $c(\tau,t)$ with respect to τ is T_m, the multipath spread of the channel, then a practical maximum of l is $L=\lceil\frac{T_m}{T_s}\rceil$, where $\lceil\cdot\rceil$ is the ceiling function (the smallest integer exceeding the argument). For frequency-nonselective fading channel $L=0$ and $h_0(t)=\int c(\tau,t)\,d\tau$. Therefore, from (2) to (5) the received signal is

$$r(t)=\sum_{l=0}^{L} h_l(t)s(t-lT_s)+\mathcal{Z}(t)\quad\begin{array}{ll} L=0 & \text{FNF} \\ L=\lceil\frac{T_m}{T_s}\rceil & \text{FSF.}\end{array} \tag{6}$$

The power spectral bandwidth of $v(t)$ is $B'=B+f_d$. Thus, $y(t)$, the output of an ideal low-pass filter with bandwidth B', is statistically sufficient to detect the data sequence (Fig. 1). Due to the flatness of the ideal low-pass filter and its bandwidth, the samples of $y(t)$, with sampling period $T_s=\frac{1}{2B'}\le\frac{1}{2B}$ such that $M=T/T_s$ becomes the

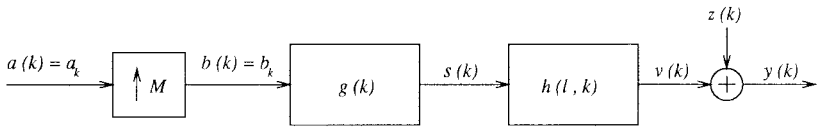


Fig. 2. The discrete model of the transmitter and the fading channel, where $\uparrow M$represents up sampling by a factor M.

minimum integer number, are given by

$$\begin{aligned} y(k) = y(t)|_{t=kT_s} &= \sum_{l=0}^{L} h_l(t)s(t-lT_s) + z(t)\Big|_{t=kT_s} \\ &= \sum_{l=0}^{L} h_l(kT_s)s(kT_s - lT_s) + z(kT_s) \\ &= h(l,k) \star s(k) + z(k) \end{aligned} \tag{7}$$

where $z(t)$, the noise component at the output of the low-pass filter, is a band-limited circularly symmetric complex Gaussian noise and $z(k) = z(t)|_{t=kT_s}$ is a circularly symmetric discrete complex Gaussian random process with variance $N_0 = \frac{1}{2}E[|z(k)|^2] = 2B'\hat{N}_0$. Also, $\star$ denotes the convolution operator, $h(l,k) = h_l(t)|_{t=kT_s}$, and $s(k) = s(t)|_{t=kT_s}$ from (1) will be

$$\begin{aligned} s(k) &= \sum_{i=0}^{\lfloor \frac{k}{M} \rfloor} a_i g(kT_s - iMT_s) \\ &= \sum_{i=0}^{k} b_i g((k-i)T_s) = b(k) \star g(k) \end{aligned} \tag{8}$$

where $g(k) = g(t)|_{t=kT_s}$ and

$$b_k = b(k) = \begin{cases} a_{\frac{k}{M}} = a\left(\frac{k}{M}\right), & \text{when } \frac{k}{M} \text{ is an integer} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Therefore, based on (7) the channel can be modeled as an L-order time-variant finite impulse response (FIR) filter. Fig. 2 shows the discrete model of the transmitter and the multipath fading channel. By uncorrelated scattering assumption for the channel with the impulse response $c(\tau, t)$, the autocorrelation function of $h(l,k)$ becomes

$$\begin{aligned} &R_h(l_1, l_2; j) \\ &\quad = \frac{1}{2}E[h(l_1, k+j)h^*(l_2, k)] \\ &\quad = \int_{-\infty}^{+\infty} \operatorname{sinc}\left(\frac{\tau - l_1T_s}{T_s}\right)\operatorname{sinc}\left(\frac{\tau - l_2T_s}{T_s}\right)R_c(\tau, jT_s)\,d\tau \\ &\qquad\qquad -\infty \le j \le \infty \quad 0 \le l_1, l_2 \le L \end{aligned} \tag{10}$$

where the autocorrelation function of $c(\tau, t)$ is

$$\begin{aligned} \phi_c(\tau, \hat{\tau}; \Delta t) &= \frac{1}{2}E[c(\tau, t+\Delta t)c^*(\hat{\tau}, t)] \\ &= R_c(\tau, \Delta t)\delta(\tau - \hat{\tau}) \\ &\quad -\infty \le \Delta t \le \infty \quad 0 \le \tau, \hat{\tau} \le \infty. \end{aligned} \tag{11}$$

As seen from (10), $h(l,k)$ for all l are wide-sense stationary random processes.

III. Maximum Likelihood Sequence Detection

The received signal is filtered by an ideal low-pass filter with bandwidth $B' = B + f_d$ and then sampled at the Nyquist sampling rate $1/T_s$ (Fig. 1). The sampler outputs $y(kT_s)$ are sufficient statistics for the detection of $A = (a_0, a_1, \cdots, a_i, \cdots, a_{I-1})$. By the assumption of equal probability for all possible input A, the MLSD is the optimal decision rule for minimizing the sequence error rate and the MLSD criterion is

$$C_{\text{MLSD}} \doteq p(Y \mid \hat{A}) = \max_{\tilde{A} \in \mathcal{A}} p(Y \mid \hat{A} = \tilde{A}) \tag{12}$$

where $\hat{A} = (\hat{a_0}, \hat{a_1}, \cdots, \hat{a_i}, \cdots, \hat{a}_{I-1})$ is the detected symbol sequence, $\mathcal{A}$ is the set of all possible transmitted symbol sequences, and $Y = (y_{K-1}, y_{K-2}, \cdots, y_k, \cdots, y_0)$ is the sequence of the output samples, where $K = MI$.[1] If the transmitter filter, whose impulse response duration is limited to L_sT_s ($L_s = \alpha M$, α is an integer), is a causal system, the output samples based on (7) and (8) are

$$\begin{aligned} y_k &= y(t)|_{t=kT_s} \\ &= \sum_{l=0}^{L} h(l,k) \sum_{n=k-l-L_s+1}^{k-l} b_n g(k-l-n) + z(k). \end{aligned} \tag{13}$$

From (12), the MLSD criterion becomes

$$\begin{aligned} &C_{\text{MLSD}} \\ &\quad = \max_{\tilde{A} \in \mathcal{A}} p(y_{K-1}, y_{K-2}, \cdots, y_k, \cdots, y_0 \mid \tilde{A}\} \\ &\quad = \max_{\tilde{A} \in \mathcal{A}} \{p(y_{K-1}, y_{K-2}, \cdots, y_{K-M} \mid y_{K-M-1}, \cdots, y_0, \tilde{A}) \\ &\qquad \times p(y_{K-M-1}, y_{K-M-2}, \cdots, \\ &\qquad y_{K-2M-1} \mid y_{K-2M-2}, \cdots, y_0, \tilde{A}) \cdots\} \\ &\quad = \max_{\tilde{A} \in \mathcal{A}} \left\{ \prod_{i=0}^{I-1} \prod_{m=0}^{M-1} p(y_{iM+m} \mid y_{iM+m-1}, \right. \\ &\qquad\qquad \left. y_{iM+m-2}, \cdots, \tilde{A}) \right\}. \end{aligned} \tag{14}$$

Let us assume that $h(l,k)$ has a finite memory length such that [3]

$$E(y_k | y_{k-1}, \cdots, y_0, \tilde{A}) = E(y_k | y_{k-1}, \cdots, y_{k-J+1}, \tilde{\boldsymbol{a}}_k). \tag{15}$$

[1] Due to the memory of the channel and the transmitter filter, the number of output samples is more than MI. However, in practice, the transmitter sends some symbols, whose length is related to the whole system memory, known by the receiver. Hence, the receiver processes only MI samples of the received signal.

Since y_k is a Gaussian random process, it can be shown that (15) leads to the relation given below[2]

$$p(y_k \mid y_{k-1}, \cdots, y_0, \tilde{A}) = p(y_k \mid y_{k-1}, \cdots, y_{k-J+1}, \tilde{\mathbf{a}}_k) \tag{16}$$

where $\tilde{\mathbf{a}}_k = [\tilde{a}_{\frac{k-L-L_s-L_R}{M}}, \cdots, \tilde{a}_{\frac{k}{M}}]$, $L_R = J-1$, and based on (9), only those elements of $\tilde{\mathbf{a}}_k$ whose subscripts are integers are included. Thus, the MLSD criterion (14) becomes

$$\begin{aligned} C_{\mathrm{MLSD}} &= \max_{\tilde{A}\in\mathcal{A}}\left\{\prod_{i=0}^{I-1}\prod_{m=0}^{M-1} p(y_{iM+m} \mid y_{iM+m-1}, \cdots, \right. \\ &\qquad \left. y_{iM+m-J+1}, \tilde{\mathbf{a}}_{iM+m})\right\} \\ &= \max_{\tilde{A}\in\mathcal{A}}\left\{\prod_{i=0}^{I-1}\prod_{m=0}^{M-1} \right. \\ &\qquad \left. \times \frac{p(y_{iM+m}, \cdots, y_{iM+m-J+1} \mid \tilde{\mathbf{a}}_{iM+m})}{p(y_{iM+m-1}, \cdots, y_{iM+m-J+1} \mid \tilde{\mathbf{a}}_{iM+m})}\right\} \\ &= \max_{\tilde{A}\in\mathcal{A}}\left\{\prod_{i=0}^{I-1}\prod_{m=0}^{M-1} \frac{\pi^{-1}(\det(R_{Y^J}))^{-1}}{(\det(R_{Y^{J-1}}))^{-1}} \right. \\ &\qquad \left. \times \frac{\exp\left(-\left(Y_{iM+m}^J\right)^{\mathcal{H}} R_{Y^J}^{-1} Y_{iM+m}^J\right)}{\exp\left(-\left(Y_{iM+m-1}^{J-1}\right)^{\mathcal{H}} R_{Y^{J-1}}^{-1} Y_{iM+m-1}^{J-1}\right)}\right\}. \end{aligned} \tag{17}$$

After using logarithms in (17), we get

$$\begin{aligned} &C_{\mathrm{MLSD}} \\ &\quad \equiv \min_{\tilde{A}\in\mathcal{A}}\left\{\sum_{i=0}^{I-1}\sum_{m=0}^{M-1} \left[\ln(\det(R_{Y^J})) - \ln(\det(R_{Y^{J-1}})) \right.\right. \\ &\qquad \left.\left. + \left(Y_{iM+m}^J\right)^{\mathcal{H}} Q(\tilde{\mathbf{a}}_{iM+m}) Y_{iM+m}^J\right]\right\} \\ &\quad \equiv \min_{\tilde{A}\in\mathcal{A}}\left\{\sum_{i=0}^{I-1}\sum_{m=0}^{M-1} B(\tilde{\mathbf{a}}_{iM+m})\right\} \end{aligned} \tag{18}$$

where $B(\cdot)$ is the branch metric, $X^{\mathcal{H}}$ denotes the conjugate transpose of X, and by defining $k = iM + m$

$$\begin{aligned} Y_k^J &= [y_k, y_{k-1}, \cdots, y_{k-J+1}]^{\mathcal{T}} \\ Y_{k-1}^{J-1} &= [y_{k-1}, y_{k-2}, \cdots, y_{k-J+1}]^{\mathcal{T}} \\ R_{Y^J} &= \frac{1}{2}E\left[Y_k^J \left(Y_k^J\right)^{\mathcal{H}} \mid \tilde{\mathbf{a}}_k\right] \\ R_{Y^{J-1}} &= \frac{1}{2}E\left[Y_{k-1}^{J-1} \left(Y_{k-1}^{J-1}\right)^{\mathcal{H}} \mid \tilde{\mathbf{a}}_k\right] \\ Q(\tilde{\mathbf{a}}_k) &= \left[R_{Y^J}^{-1}\right] - \begin{bmatrix} 0 & \vdots & 0_{1\times(J-1)} \\ \cdots\cdots & \cdots & \cdots\cdots \\ 0_{(J-1)\times 1} & \vdots & R_{Y^{J-1}}^{-1} \end{bmatrix} \end{aligned}$$

[2]Conditions in (15) and (16) can be interpreted as wide-sense and strict-sense Markov properties. It can be shown that if a Gaussian sequence is wide-sense Markov (15), it is also strict-sense Markov (16) [25].

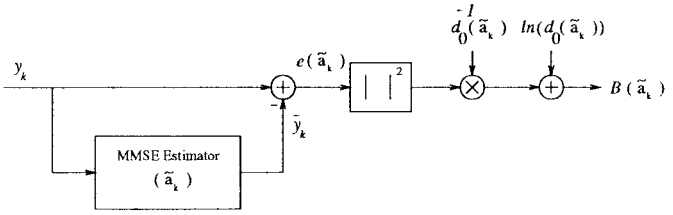


Fig. 3. The block diagram of branch metric evaluation.

where $X^{\mathcal{T}}$ denotes the transpose of X and $0_{i\times j}$ is a $i\times j$ zero matrix. Using (13) the ijth element of R_{Y^J} is

$$\begin{aligned} &R_Y(i,j) \\ &\quad = \frac{1}{2}E[y_{k-i}y_{k-j}^* \mid \tilde{\mathbf{a}}_k] \\ &\quad = \frac{1}{2}E\left[\left(\sum_{l_1=0}^{L} h(l_1, k-i)s(k-i-l_1) + z(k-i)\right)\right. \\ &\qquad \left.\times\left(\sum_{l_2=0}^{L} h^*(l_2, k-j)s^*(k-j-l_2) + z^*(k-j)\right)\right] \\ &\quad = \sum_{l_1=0}^{L}\sum_{l_2=0}^{L} R_h(l_1, l_2; j-i)s(k-i-l_1) \\ &\qquad \times s^*(k-j-l_2) + N_0\delta(j-i). \end{aligned} \tag{19}$$

From (10), it follows that $R_h(l_1,l_2;j) = R_h(l_2,l_1;j)$, $R_h(l_1,l_2;j) = R_h^*(l_1,l_2;-j)$, and $R_Y(i,j) = R_Y^*(j,i)$. Therefore, R_{Y^J} is a symmetric positive definite matrix, and it can be factorized as $R_{Y^J} = L_J(\tilde{\mathbf{a}}_k)D_J(\tilde{\mathbf{a}}_k)L_J^{\mathcal{H}}(\tilde{\mathbf{a}}_k)$ [26], where $L_J(\tilde{\mathbf{a}}_k)$ is a unit upper triangular matrix and $D_J(\tilde{\mathbf{a}}_k)$ is a diagonal matrix, $D_J(\tilde{\mathbf{a}}_k) = \mathrm{diag}(d_0(\tilde{\mathbf{a}}_k), d_1(\tilde{\mathbf{a}}_k), \cdots, d_{J-1}(\tilde{\mathbf{a}}_k))$. Then the inverse of R_{Y^J} is given as

$$R_{Y^J}^{-1} = \Gamma_J^{\mathcal{H}}(\tilde{\mathbf{a}}_k)D_J^{-1}(\tilde{\mathbf{a}}_k)\Gamma_J(\tilde{\mathbf{a}}_k) \tag{20}$$

where $\Gamma_J(\tilde{\mathbf{a}}_k) = L_J^{-1}(\tilde{\mathbf{a}}_k)$. In the same way, the inverse of $R_{Y^{J-1}}$ can be calculated. After some manipulations (Appendix A), the branch metric $B(\tilde{\mathbf{a}}_k)$ in (18) becomes

$$B(\tilde{\mathbf{a}}_k) = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{\left|\sum_{j=0}^{J-1} y_{k-j}\gamma_{j,0}(\tilde{\mathbf{a}}_k)\right|^2}{d_0(\tilde{\mathbf{a}}_k)} \tag{21}$$

where $\gamma_{i,j}(\tilde{\mathbf{a}}_k)$ is the ijth element of the $\Gamma_J(\tilde{\mathbf{a}}_k)$ matrix. It can be shown that the factorization of $R_{Y^J}^{-1}$ based on (20) is equivalent to a Gram–Schmidt orthogonalization of Y_k^J [27]. $L_J(\tilde{\mathbf{a}}_k)$, and $\Gamma_J(\tilde{\mathbf{a}}_k)$ are called innovations and whitening filters, respectively [26, p. 207], such that by feeding the colored random process y_{k-j} to the FIR filter with coefficients $\gamma_{j,0}(\tilde{\mathbf{a}}_k)$ for $j = 0, \cdots, J-1$, the output becomes a whitened random process. The branch metrics in (21) can also be written as

$$B(\tilde{\mathbf{a}}_k) = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{|e(\tilde{\mathbf{a}}_k)|^2}{d_0(\tilde{\mathbf{a}}_k)} \tag{22}$$

where, as shown in Fig. 3, $e(\tilde{\mathbf{a}}_k) = y_k - \tilde{y}_k$ is the prediction error and $\tilde{y}_k = -\sum_{j=1}^{J-1} y_{k-j}\gamma_{j,0}(\tilde{\mathbf{a}}_k)$ is the causal minimum-mean-square error (MMSE) estimate of y_k based on the assumption of sending $\tilde{\mathbf{a}}_k$ and given the previous samples

$y_{k-1}, \cdots, y_{k-J+1}$ [26]. $d_0(\tilde{\mathbf{a}}_k)$ is the variance of the error $e(\tilde{\mathbf{a}}_k)$, and, thus, the branch metric can be interpreted as the normalized prediction error with a bias equal to $\ln(d_0(\tilde{\mathbf{a}}_k))$.

Using the dynamic programming algorithm based on (18), the receiver detects the information sequence A. When using q-ary symbols, the number of states in the trellis diagram is $q^{\vartheta-1}$, where $\vartheta = \lceil \frac{L+L_s+L_R}{M} \rceil$ and the MLSD receiver for the detection of each symbol between i and $i+1$ stages in the trellis diagram calculates Mq^{ϑ} branch metrics from (18) and finds the survivor paths or minimum cost paths which terminate at each state in the trellis diagram.

IV. Adaptive State Allocation Algorithm

The MLSD receiver based on (18) is the optimal receiver for multipath fading channels. However, the major challenge with this nonlinear receiver is its computational complexity. The complexity of the MLSD receiver depends on calculating the branch metrics of paths terminating at each state, finding the minimum of these metrics and keeping the survivor path for each state. While the first two factors are related to computational complexity, the last factor is related to necessary storage. Of these three factors, calculating the branch metrics is the major contribution to the complexity, and sometimes decreasing the complexity of this factor also decreases that of the other factors. In this section, we consider the computational complexity associated with the calculation of the branch metrics.

A. Structure

In general, the number of multiplications per symbol, N_m, in the MLSD receiver is

$$N_m = qM_bN_s \tag{23}$$

where q is the size of the signal set (q branches emanate from each state), M_b is the number of multiplications used in the calculation of each branch metric, and N_s is the number of states in the trellis. $N_s = q^{\vartheta-1}$, where ϑ depends on the parameters of the channel and the shaping filter in the transmitter. Reducing the computational complexity can be achieved by decreasing any one of parameters in (23) or a combination of them.

The trellis states at stage i are denoted as $S_i = (\tilde{a}_{i+2-\vartheta}, \tilde{a}_{i+3-\vartheta}, \cdots, \tilde{a}_i)$, where the number of states depends on the size of the signal set that every symbol $\tilde{a}_{i-j}$ can select in stage i. The maximum size is q and the maximum number of states is $q^{\vartheta-1}$. If q_{ij} is the size of signal set that symbol $\tilde{a}_{i-j}$ can select at stage i, the total number of states in the trellis will be

$$N_{s_i} = \prod_{j=1}^{\vartheta-1} q_{ij}. \tag{24}$$

Three different ways exist for reducing the number of states at stage i. The first way is to fuse or combine some states, which have the common latest symbol set, to one state. For example, in a binary alphabet, two states $(\tilde{a}_{i-2}, \tilde{a}_{i-1}, \tilde{a}_i) \equiv 100$ and $(\tilde{a}_{i-2}, \tilde{a}_{i-1}, \tilde{a}_i) \equiv 000$ can be fused to state $(\tilde{a}_{i-1}, \tilde{a}_i) \equiv 00$, since the latest symbol set $(\tilde{a}_{i-1}, \tilde{a}_i)$ of both is common. However, the two states (101) and (000) cannot be fused. This is similar to the partitioning proposed in [16], but in contrast to the method in [16], where only a fixed or static partitioning is used for all stages, in the present work we use adaptive partitioning.

The second way to reduce the number of states is to select a subset with N_{b_i} states where $N_{b_i} \leq N_{s_i}$. This method is similar to the M algorithm [15]. However, the number of selected states is not necessarily the same at each stage. The third way is the combination of the above two methods at the same time. Not only is the number of states reduced by fusion, but also a subset of fused states is chosen. This method is suitable for time-variant channels like fading channels, where the signal-to-noise ratio (SNR) is time variant. When the channel goes to fade, the SNR is low and it is very likely that an error will occur. When the channel is out of fade the SNR is high and the error rate is very much less than the average error rate. Therefore, using the same computational complexity for all SNR situations is not efficient. Fig. 4 shows the structure of a trellis diagram based on this state reduction method called the ASA algorithm. In this algorithm, the number of states is reduced by fusion when the channel is out of fade and the number of states is increased by diffusion when the channel goes to fade. Also, at each stage, only the more likely correct states are chosen for extension to the next stage.

The trellis diagram of ASA algorithm in Fig. 4 is an example for $q = 2$. In stage i, since the short-term power of the received signal is very low, the receiver considers all the states as the more likely correct states. The short-term power of the received signal in stage $i+1$ is also low, however, only states $(000), (110), (001), (011)$, and (111) are recognized as the more likely correct states and extended to the next stage $i+2$. In a similar way in stage $i+2$, only $(010), (110), (001), (101)$, and (111) are continued to the next stage. The short-term power of the received signal is high in stage $i+3$, therefore, the states having the same latest symbol set are fused. In stage $i+4$, only (10) and (01) states are pointed out as the more likely correct states and extended to the subsequent stage. Since the short-term power of the received signal at stage $i+5$ is low, the states are diffused and the number of symbols for calculating the branch metrics is increased as well.

The other parameter, which is important in the computational complexity, is M_b in (23). Conventionally, the number of multiplications for calculating branch metrics is proportional to ϑ or more precisely to the number of elements in $S_i = (\tilde{a}_{i+2-\vartheta}, \tilde{a}_{i+3-\vartheta}, \cdots, \tilde{a}_i)$. By fusion, the size of the signal set for some symbols is only one and the number of multiplications can be reduced or it remains the same. For example, when $\hat{S}_i = (\tilde{a}_{i+2-\hat{\vartheta}}, \cdots, \tilde{a}_i)$ is fused to $S_i = (\tilde{a}_{i+2-\vartheta}, \cdots, \tilde{a}_i)$, the many alternatives for the partial sequence of $\hat{S}_i$, $(\tilde{a}_{i+2-\hat{\vartheta}}, \cdots, \tilde{a}_{i+1-\vartheta})$, is replaced by the sequence of previously detected symbols. If sample vectors $Y_k^{\hat{J}}$ and Y_k^J correspond to symbol vectors $\hat{S}_i$ and S_i, respectively, in the fusion situation, the branch metrics which originate from each state can be calculated in association with $Y_k^{\hat{J}}$ or Y_k^J. Generally, Y_k^J is selected for fusion states; but if $Y_k^{\hat{J}}$ is chosen the symbols which have only one option in the signal set are considered as decision symbols in calculating

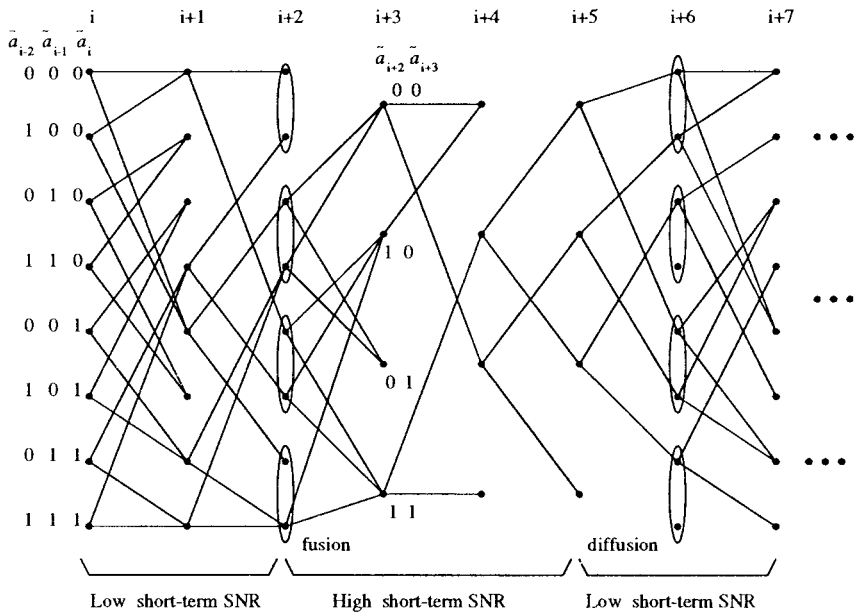


Fig. 4. Structure of the trellis diagram in ASA algorithm with $q = 2$.

the branch metrics. This method is called decision feedback sequence estimation (DFSE) [16]. In the trellis structure of the ASA algorithm, the branch metrics can be calculated with or without decision symbols. If decision symbols are used we call it ASA-DF. We define $Y_k^{\hat{J}} = Y_k^{J+F} = [Y_k^J \vdots Y_k^F]$, where $Y_k^F = [y_{k-J}, \cdots, y_{k-J-F+1}]$ and F is the number of previous samples corresponding to decisions. Similar to (21), the new branch metrics for ASA-DF become

$$B(\tilde{\mathbf{a}}_k) = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{\left|\sum_{j=0}^{J+F-1} y_{k-j}\gamma_{j,0}(\tilde{\mathbf{a}}_k)\right|^2}{d_0(\tilde{\mathbf{a}}_k)}. \quad (25)$$

In fading channels, due to error propagation effects, there is a limit to performance improvement obtainable by increasing F. This restricts the decrease in the computational complexity which can result from decision feedback. The combination of ASA and DF is an alternative way for having a good tradeoff between performance and complexity.

B. Strategy

The main strategy of the ASA algorithm is to keep the performance of the ASA close to that of MLSD with a reduction of computational complexity associated with the trellis structure in Fig. 4. It is well known that the error probability in the MLSD for ISI channels depends on the minimum distance of the shortest error event ($d_{\min}$) in the trellis structure [4]. In contrast to ISI channels, $d_{\min}$ is not constant in fading channels and depends on the state of the channel. When the channel goes to fade, $d_{\min}$ is smaller than when the channel is out of fade. The MLSD performance for fading channels is mostly determined by the channel when it is in fade or the received signal power is lower than the average signal power. Therefore, the strategy of the ASA algorithm should be designed to keep the minimum distance error event of the ASA close to that of the MLSD, especially in a low short-term SNR situation.

Fig. 3 shows how the branch metrics (for correct or incorrect path) are calculated based on MMSE estimation. The algorithm calculates the branch metrics in the trellis diagram for all the possibilities of sending $\mathbf{a}_k$, however, only one of them has been sent by the transmitter and we call this branch metric as the correct branch metric (or correct path). The other branch metrics are called incorrect branch metrics. We define $E[e_u(\tilde{\mathbf{a}}_k)] = E[e(\tilde{\mathbf{a}}_k) \mid \tilde{\mathbf{a}}_k \neq \mathbf{a}_k]$ and $E[e_c(\tilde{\mathbf{a}}_k)] = E[e(\tilde{\mathbf{a}}_k) \mid \tilde{\mathbf{a}}_k = \mathbf{a}_k]$ as the expectation of the estimation errors corresponding to the incorrect and the correct branch metrics, respectively. Also, $\tilde{\mathbf{a}}_k$ and $\mathbf{a}_k$ are the data in the branches and the actual transmitted data, respectively. Since the channel is fading, each sample y_{k-j} is a complex-valued random variable even if the transmitted signal $s(t)$ is known and the additive Gaussian noise is zero. Therefore, $e_u(\tilde{\mathbf{a}}_k)$ and $e_c(\tilde{\mathbf{a}}_k)$ are complex-valued Gaussian random variables with zero mean. The expectation of the incorrect branch metrics are

$$\begin{aligned} &E(B(\tilde{\mathbf{a}}_k) \mid \tilde{\mathbf{a}}_k \neq \mathbf{a}_k) \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{E[|e(\tilde{\mathbf{a}}_k)|^2 \mid \tilde{\mathbf{a}}_k \neq \mathbf{a}_k]}{d_0(\tilde{\mathbf{a}}_k)} \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{E[|e_u(\tilde{\mathbf{a}}_k)|^2]}{d_0(\tilde{\mathbf{a}}_k)} \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{\varepsilon_u(\tilde{\mathbf{a}}_k)}{d_0(\tilde{\mathbf{a}}_k)} \end{aligned} \quad (26)$$

where $\varepsilon_u(\tilde{\mathbf{a}}_k) = E[|e_u(\tilde{\mathbf{a}}_k)|^2]$ and the expectation of the correct branch metric is

$$\begin{aligned} &E(B(\tilde{\mathbf{a}}_k) \mid \tilde{\mathbf{a}}_k = \mathbf{a}_k) \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{E[|e(\tilde{\mathbf{a}}_k)|^2 \mid \tilde{\mathbf{a}}_k = \mathbf{a}_k]}{d_0(\tilde{\mathbf{a}}_k)} \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{E[|e_c(\tilde{\mathbf{a}}_k)|^2]}{d_0(\tilde{\mathbf{a}}_k)} \\ &\quad = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{\varepsilon_c(\tilde{\mathbf{a}}_k)}{d_0(\tilde{\mathbf{a}}_k)}. \end{aligned} \quad (27)$$

It can be shown [26, p. 504] that

$$\varepsilon_c(\tilde{\mathbf{a}}_k) = E[|e_c(\tilde{\mathbf{a}}_k)|^2] = d_0(\tilde{\mathbf{a}}_k). \tag{28}$$

According to correct path assumption, $\varepsilon_c(\tilde{\mathbf{a}}_k)$ is the expectation of the true minimum mean square error. Thus, $\varepsilon_u(\tilde{\mathbf{a}}_k) \geq \varepsilon_c(\tilde{\mathbf{a}}_k)$. The probability density function (pdf) of $|e_c(\tilde{\mathbf{a}}_k)|^2$ and $|e_u(\tilde{\mathbf{a}}_k)|^2$ are chi-square with two degrees of freedom [26, p. 117]. If we define new random variables $\eta_c = \frac{|e_c(\tilde{\mathbf{a}}_k)|^2}{d_0(\tilde{\mathbf{a}}_k)}$ and $\eta_u = \frac{|e_u(\tilde{\mathbf{a}}_k)|^2}{d_0(\tilde{\mathbf{a}}_k)}$, the means of η_c and η_u are $\bar{\eta_c} = \frac{\varepsilon_c(\tilde{\mathbf{a}}_k)}{d_0(\tilde{\mathbf{a}}_k)} = 1$ and $\bar{\eta_u} = \frac{\varepsilon_u(\tilde{\mathbf{a}}_k)}{d_0(\tilde{\mathbf{a}}_k)} \geq 1$ and also the pdf's of η_c and η_u are

$$f_c(x) = \exp(-x)u(x) \tag{29}$$

$$f_u(v) = \frac{1}{\bar{\eta}_u}\exp\left(\frac{-v}{\bar{\eta}_u}\right)u(v) \tag{30}$$

where $u(\cdot)$ is the unit step function. Therefore, the pdf's of branch metrics for correct and incorrect branch metrics are also chi-square with two degrees of freedom that are biased with a constant $\ln(d_0(\tilde{\mathbf{a}}_k))$.

As seen in Section III, the derivation of the structure of the MLSD receiver for fading channels was based on obtaining the shortest path in the trellis diagram. Generally, in the dynamic programming algorithm all branch metrics, which originate from each state, are calculated. In other words, all states have the same potential to be the correct state. However, the performance will not be affected by cancelling those states for which the probabilities of being the correct state are very small. The strategy of the ASA algorithm is to select only the more likely correct states for extension to the next stage such that with high probability the correct state is among the more likely correct states. If S_j is defined as the set of all possible states in the trellis diagram at stage j and $C_j(\varsigma_\kappa)$ is the cost of survivor path terminating to state ς_κ ($\varsigma_\kappa \in S_j$) at stage j, the minimum cost at this stage is defined by

$$\begin{aligned} C_{\min} &= \min_{\tilde{A}\in\mathcal{A}}\left\{\sum_{i=0}^{j}\sum_{m=0}^{M-1} B(\tilde{\mathbf{a}}_{iM+m})\right\} \\ &= \min_{S_j}\{C_j(\varsigma_\kappa)\}. \end{aligned} \tag{31}$$

In each stage, the ASA algorithm finds the minimum cost state, which is the final state in the survivor path which has the minimum cost, and selects states whose costs are

$$C_j(\varsigma_\kappa) \leq C_{\min} + T_h, \qquad \text{for all } \varsigma_\kappa \in S_j. \tag{32}$$

We call those states satisfying (32) as the "more likely correct states," and only the survivor paths from the more likely correct states extend to the next stage. T_h in (32) should be selected such that the correct state is among the more likely correct states with high probability; this will be discussed now.

If we define a new random variable $\zeta = \eta_m - \eta_c$, where η_m is η_u which has the minimum expectation over all incorrect branch metrics ($\bar{\eta}_m \leq \min\{\bar{\eta}_u\}$), the pdf of ζ by independency assumption between η_c and η_m will be

$$f_\zeta(x) = \begin{cases} \frac{1}{1+\bar{\eta}_m}e^{x}, & x \leq 0 \\ \frac{1}{1+\bar{\eta}_m}e^{\frac{-x}{\bar{\eta}_m}}, & x \geq 0. \end{cases} \tag{33}$$

If P_0 is defined as

$$P_0 = \int_{-T_h}^{+\infty} f_\zeta(x)\,dx, \qquad T_h \geq 0 \tag{34}$$

the tradeoff between performance and complexity is controlled by P_0. Also, the probability of being the correct state among the more likely correct states relates to the value of P_0. When $\bar{\zeta} = E(\zeta)$ is close to zero (always $\bar{\zeta} \geq 0$), the probability of the minimum cost state being the correct state decreases. In this situation, increasing T_h (or P_0) is equivalent to increasing the probability of the correct state being among the selected states or to improving the performance. However, when $\bar{\zeta}$ is very much larger than zero, the probability of the minimum state being the correct state increases. In this situation increasing T_h (or P_0) is equivalent to increasing the probability of adding more computational complexity without performance gain. Therefore, T_h is a function of the distance between $f_c(x)$ and $f_m(x)$, where $f_m(x)$ is the pdf of η_m. In general, there are different definitions for the distance between two pdf's [28]. Here, we use the Chernoff distance d_ℓ between $f_m(x)$ and $f_c(x)$, where

$$d_\ell = -\ln\left[\int f_c^\ell(x) f_m^{1-\ell}(x)\,dx\right], \qquad 0 \leq \ell \leq 1. \tag{35}$$

Associated with (36), the threshold T_h will be defined as

$$T_h = cd_\ell + \bar{\eta}_c \tag{36}$$

where c is a constant value. ℓ and c in (35) and (36) should be set by selecting P_0 for each channel based on a tradeoff between performance and complexity. We choose P_0 close enough to one at two high SNR values such that the performance of ASA and MLSD become close and obtain two corresponding T_h from (34) [note that $f_\zeta(x)$ and hence T_h depend on SNR]. Then c and ℓ can be set based on (35) and (36), and for these fixed c and ℓ values, T_h can be obtained from (36) for different values of SNR.

The threshold T_h in (36) was derived based on the long-term SNR. In fading channels, however, the power of the received signal is changing with time. Therefore, the power of estimation error is changing with time and the threshold should also be changing with time. Therefore, we scale the threshold T_h in (36) based on short- and long-term averages of received signal power

$$T_h = (cd_\ell + \bar{\eta}_c)\frac{P_s}{P} \tag{37}$$

where $P = \frac{1}{2}E[|y_k|^2]$ and $P_s = \frac{1}{2J}\sum_{j=0}^{J-1}|y_{k-j}|^2$ are the long- and the short-term averages of received signal power, respectively. Now we consider the strategy of fusing and diffusing states in the ASA algorithm. As mentioned before, the threshold is chosen such that the correct state is among the selected states with high probability. Thus, when the number of the more likely correct states is one, it means that the status of channel is good, but when the number of the more likely correct states is more than one, it means that the status of the channel is not good enough and error events may occur. By diffusing the states, the difference of estimation error between correct and incorrect states increases and also the

number of branch metrics in the error event increases. Then the probability of the error event will decrease. By fusing states, however, this situation will be reversed. Therefore, the strategy of fusing and diffusing states is based on the number of the more likely correct states.

In the ASA algorithm, when the number of the more likely correct states is decreasing, the structure of the trellis does not change until the number of the more likely correct states becomes one. In this situation, the states of the trellis diagram are fused. However, when the number of the more likely correct states is increasing or remains constant (more than one stage) and expectation of the minimum distance error event $\bar{d}_{\min}$ for the next step of diffusing states is more than that of the current step, the states of the trellis diagram are diffused. Similar to T_h, the expectation of the minimum distance error event is normalized based on the short- and long-term averages of the received signal power

$$\bar{d}_{\min} = \jmath\bar{\zeta}\frac{P_s}{P} \tag{38}$$

where $\jmath$ is the shortest length of error event in trellis diagram. Now we describe the steps of the ASA algorithm. Given the autocorrelation function of the channel impulse response $g(t)$, the minimum and maximum of J ($J_{\min}$ and $J_{\max}$), c, d_ℓ, and $\bar{\zeta}$, the ASA algorithm follows these steps.

1) Receive the vector Y_k^J and calculate branch metrics which originate from the more likely correct states by using (21).
2) Find $C_{\min}$ from (31).
3) Calculate P_s from Y_k^J and find T_h and $\bar{d}_{\min}$ from (37) and (38), respectively [P is known from the autocorrelation function of the channel impulse response and $g(t)$].
4) Find the more likely correct states from (32) and determine N_{b_i}, the number of the more likely correct states in stage i.
5) If $N_{b_i} = N_{b_{i-1}} = 1$, fuse the states (select new J) and go to step 1). Else if $N_{b_i} \geq N_{b_{i-1}} \geq N_{b_{i-2}}$, $N_{b_{i-1}} \neq 1$ and $\bar{d}_{\min}$ (diffusion) $> \bar{d}_{\min}$(present) diffuse the states and go to step 1). Else go to step 1).

The ASA-DF algorithm is similar to ASA, except that the branch metrics will be calculated from (24) and short-term power is obtained from

$$P_s = \frac{1}{2(J+F)} \sum_{j=0}^{J+F-1} |y_{k-j}|^2. \tag{39}$$

Although in this section we proposed a specific method for obtaining T_h and $\bar{d}_{\min}$ for fusion and diffusion of states in fading channels, one can also propose other methods of optimization based on the different levels of channel knowledge.

C. Implementation

The number of states and branch metrics associated with Fig. 4 are time variant in the ASA algorithm. Thus, the computational complexity of ASA algorithm is also time variant and is much lower than that of the regular MLSD only in the mean sense. From the point of real-time implementation of the ASA algorithm, the software or hardware should be able to handle the ASA algorithm with the highest possible computational complexity which equals that of the regular MLSD. It means that the full computational complexity of software or hardware is used in the ASA algorithm only when the short-term SNR is very low. Although the time-variant computational complexity of the ASA algorithm decreases the power consumption, this is not a strong reason for using the ASA algorithm instead of the regular MLSD.

For making the computational complexity of software or hardware constant, we propose to use a buffering shift register after the sampler in Fig. 1, where the samples of the received signal are held in the shift register and then are used by the digital processor. Although the input rate to the shift register is $\frac{1}{T_s}$ and is constant, the output rate from the shift register is time variant and depends on the short-term SNR. When the short-term SNR increases (decreases), time needed to process the samples decreases (increases) and hence the output rate of the shift register increases (decreases). Therefore, from the viewpoint of software or hardware, the computational complexity is constant and only the rate of the signal received by the digital processor is time variant. In this situation, the software or hardware can be designed to handle the mean value of the time-variant computational complexity of the ASA algorithm.

It is clear that the length of the queue in the shift register is time variant. The shift-register length should be selected based on the statistical characteristics of the queue length in shift register which depends on the channel coherence time, channel multipath spread and sampling rate. The design of the shift register is beyond the scope of this paper and we assume that the shift register does not overflow.

V. Diversity Reception

Errors in the fading channels occur in burst mode and the use of diversity is a well-known method for improving the performance. The MLSD criterion for the reception of the same signal from D independently fading channels is

$$C_{\text{MLSD}} \doteq \max_{\tilde{A}\in\mathcal{A}} p(Y^{(1)}, Y^{(2)}, \cdots, Y^{(D)} \mid \hat{A} = \tilde{A}) \tag{40}$$

where $Y^{(d)} = \{y_{K-1}^{(d)}, y_{K-2}^{(d)}, \cdots, y_0^{(d)}\}$ is the vector received signal from channel d and $y_k^{(d)}$ is

$$y_k^{(d)} = \sum_{l=0}^{L} h^{(d)}(l,k) \sum_{n=k-l-L_s+1}^{k-l} b_n g((k-l-n)) + z^{(d)}(k), \quad d = 1, 2, \cdots, D. \tag{41}$$

Based on the independence of D channels and the results of Section III, we get

$$\begin{aligned} C_{\text{MLSD}} &\equiv \min_{\tilde{A}\in\mathcal{A}} \left\{ \sum_{i=0}^{I-1} \sum_{m=0}^{M-1} \sum_{d=1}^{D} B^{(d)}(\tilde{\mathbf{a}}_{iM+m}) \right\} \\ &= \min_{\tilde{A}\in\mathcal{A}} \left\{ \sum_{i=0}^{I-1} \sum_{m=0}^{M-1} \mathbf{B}(\tilde{\mathbf{a}}_{iM+m}) \right\} \end{aligned} \tag{42}$$

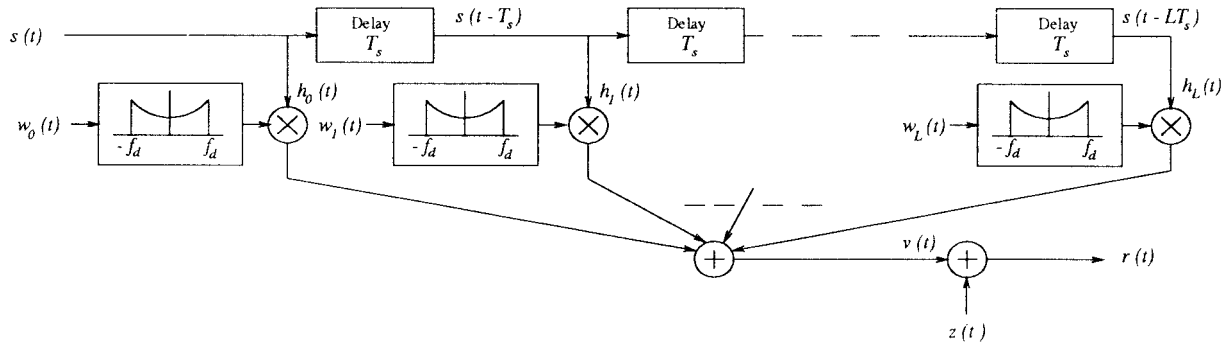


Fig. 5. The baseband tapped delay line model of the multipath fading channel.

where the branch metric $\mathbf{B}(\tilde{\mathbf{a}}_k)$ is given by

$$\begin{aligned}\mathbf{B}(\tilde{\mathbf{a}}_k) &= \sum_{d=1}^{D} B^{(d)}(\tilde{\mathbf{a}}_k) = \sum_{d=1}^{D} \ln\big(d_0^{(d)}(\tilde{\mathbf{a}}_k)\big) \\ &\quad + \frac{\left|\sum_{j=0}^{J-1} y_{k-j}^{(d)} \gamma_{j,0}^{(d)}(\tilde{\mathbf{a}}_k)\right|^2}{d_0^{(d)}(\tilde{\mathbf{a}}_k)} \\ &= \sum_{d=1}^{D} \ln\big(d_0^{(d)}(\tilde{\mathbf{a}}_k)\big) + \frac{\left|e^{(d)}(\tilde{\mathbf{a}}_k)\right|^2}{d_0^{(d)}(\tilde{\mathbf{a}}_k)} \\ &= \sum_{d=1}^{D} \ln\big(d_0^{(d)}(\tilde{\mathbf{a}}_k)\big) + \sum_{d=1}^{D} \eta^{(d)}. \end{aligned} \tag{43}$$

If the random variable $\eta^{(d)}$ is called $\eta_c^{(d)}$ and $\eta_u^{(d)}$ for the correct and incorrect paths (branch metrics), respectively, we define

$$\eta_c = \sum_{d=1}^{D} \eta_c^{(d)} \quad \text{correct path} \tag{44}$$

$$\eta_u = \sum_{d=1}^{D} \eta_u^{(d)} \quad \text{incorrect path.} \tag{45}$$

As mentioned in Section IV-B, the pdf's of $\eta_c^{(d)}$ and $\eta_u^{(d)}$ are chi-square with two degrees of freedom. Therefore, the pdf's of η_c and η_u will also be chi-square with $2D$ degrees of freedom

$$\begin{aligned} f_c(\eta_c) &= f_c\big(\eta_c^{(1)}\big) * f_c\big(\eta_c^{(2)}\big) * \cdots * f_c\big(\eta_c^{(D)}\big) \\ &= \frac{\eta_c^{D-1}}{(D-1)!} e^{-\eta_c} u(\eta_c) \end{aligned} \tag{46}$$

$$\begin{aligned} f_u(\eta_u) &= f_u\big(\eta_u^{(1)}\big) * f_u\big(\eta_u^{(2)}\big) * \cdots * f_u\big(\eta_u^{(D)}\big) \\ &= \sum_{d=1}^{D} k^{(d)} e^{\frac{-\eta_u}{\bar{\eta}_u^{(d)}}} u(\eta_u) \end{aligned} \tag{47}$$

where $k^{(d)}$ for $d = 1, 2, \cdots, D$ are functions of $\bar{\eta}_u^{(d)}$. The threshold T_h is also defined similar to (36). Although the dynamic range of the changing T_h for flat fading is large when $D = 1$, it is very small in the selective fading channel, especially when $D \gg 1$. In general, diversity reception decreases the dynamic range of the time-variant SNR in fading channels. Thus, the dynamic range of the computational complexity of the ASA algorithm is decreased as well.

VI. Simulation Results and Comparisons

To evaluate the bit-error rate performance of the MLSD, ASA, and ASA-DF, computer simulations were done for two different fading channels with $f_dT = 0.1$:

1) flat fading channel;
2) exponential channel with autocorrelation function

$$\phi_c(\tau, \hat{\tau}; \Delta t) = \sum_{l} \exp(-b\tau) J_0(2\pi f_d \Delta t) \delta(\tau - \hat{\tau}) \delta(\tau - lT_s),$$
$$0 \le \tau, \hat{\tau} \le \infty \quad -\infty \le \Delta t \le \infty \tag{48}$$

where J_0 is the zero-order Bessel function and f_d is the maximum Doppler frequency. The delay rate in (48) was chosen as $b^{-1} = 2T_s$. We simulated the exponential channel with five paths associated with (6) or $L = 4$. According to (10), the autocorrelation function of the discrete model of channel, $R_h(l_1, l_2; j)$ will be

$$\begin{aligned} R_h(l_1, l_2; j) &= R_h(l_1, l_2; \Delta t)|_{\Delta t = jT_s} \\ &= \sum_{l_1=0}^{L} \exp(-bl_1T_s) J_0(2\pi f_d j T_s) \delta(l_1 - l_2), \\ &\quad 0 \le l_1, l_2 \le L \quad -\infty \le j \le \infty. \end{aligned} \tag{49}$$

As seen for this model of the channel, the uncorrelated scattering assumption is also valid for the discrete model of channel. Based on (6), a tapped delay line model of channel is shown in Fig. 5, where $w_l(t)$ for $l = 0, \cdots, L$ and $\mathcal{Z}(t)$ are mutually independent zero mean complex white Gaussian random processes and $h_l(t)$ can be obtained by filtering $w_l(t)$ through a filter whose transfer function is proportional to the square root of the power spectral density of $h_l(t)$.

In channel simulation, the memory of $h_l(k)$ with respect to k is $L_R = 8$. In order to limit the complexity of computer simulation, we consider the memory duration $J = 3$, 5, and 7 in (16) and we call them MLSD-2, MLSD-3, and MLSD-4, respectively. In the ASA algorithm, the minimum and maximum of J are $J_{\min} = 3$ and $J_{\max} = 7$. By selecting $P_0 = 0.999$ in (34), such that with high probability the correct state is selected as one of the more likely correct states at

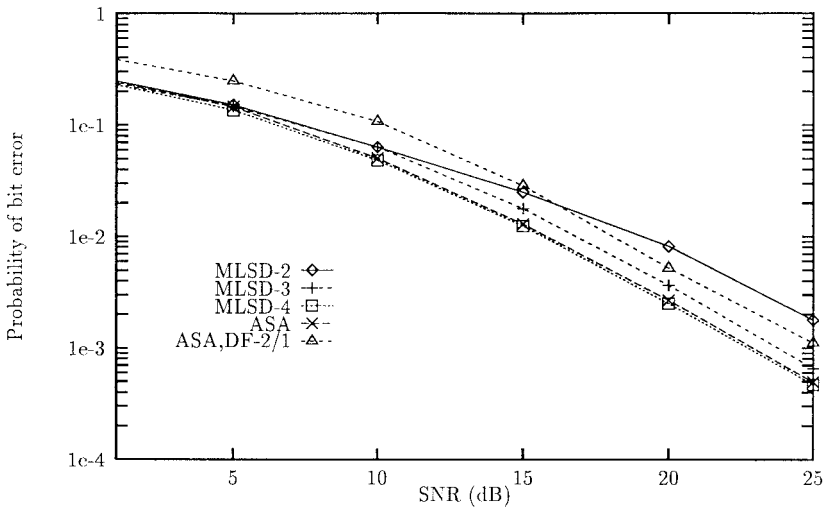


Fig. 6. Bit-error rate performance of flat fading for MLSD-2, MLSD-3, MLSD-4, ASA and ASA-DF-1/2 with $D = 1$ and for $f_dT = 0.1$.

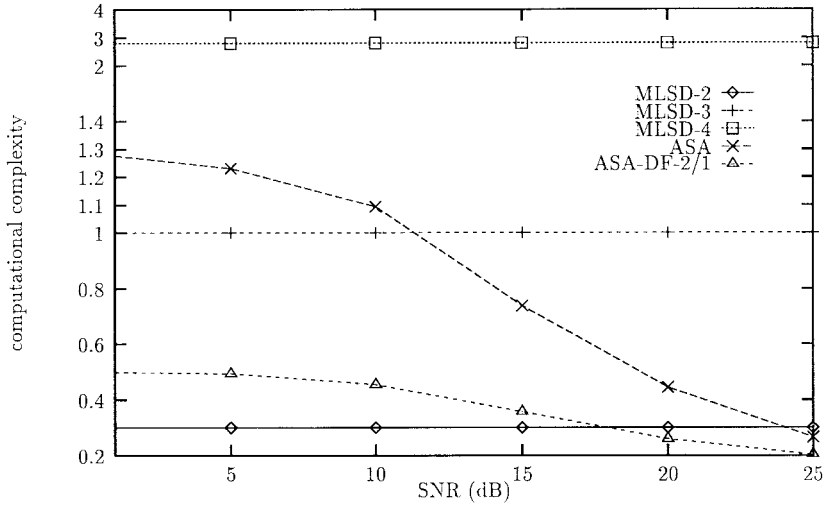


Fig. 7. Computational complexity of flat fading for MLSD-2, MLSD-3, MLSD-4, ASA, and ASA-DF-1/2 with $D = 1$ for $f_dT = 0.1$. The computationalcomplexity of all systems is normalized to that ofMLSD-3 with $D = 1$.

high SNR's, ℓ for all channel models and SNR's is obtained as $\ell = 0.9$. However, the value of c is different for each channel model.

Since the main purpose of the ASA-DF is to decrease the number of states without performance degradation by using the decision symbols to estimate the received symbol accurately, we select a fixed minimum number of states or $J = J_{\min} = 3$ in the ASA-DF algorithm. The systems are called ASA-DF2/1 and ASA-DF2/2 when in association with (25) $J = 3$, $F = 2$ and $J = 3$, $F = 4$, respectively. Meanwhile, the signaling scheme is orthogonal or in terms of (1), each information symbol a_i belongs to the set $\{1, j\}$ $(j = \sqrt{-1})$. The impulse response of the transmitter filter is a raised-cosine pulse

$$g(t) = \text{sinc}\left(\frac{t}{T}\right)\left(\frac{\cos(\nu\pi t/T)}{1-(2\nu t/T)^2}\right) \tag{50}$$

where the symbol duration $T = 1$ and ν is the roll of factor. The bandwidth of the signal, therefore, is $B = \frac{1+\nu}{2T}$ and the sampling rate is $T_s \leq \frac{T}{1+\nu+2f_dT}$. Based on different f_d, ν was selected such that $\frac{T}{1+\nu+2f_dT} \geq T/2$. In the simulation, we chose $T_s = T/2$, $M = 2$. For reducing the amount of algorithm computation, especially for selective fading channels, the duration of the $g(t)$ was truncated to $4T_s$ or $L_s = 4$ (13).

Figs. 6 and 10 show the performances of MLSD-2, MLSD-3, MLSD-4, ASA, and ASA-DF-2/1 for the flat fading channel with $D = 1$ and $D = 2$, respectively. As can be seen, the performances of ASA are very close to that of MLSD-4 both in $D = 1$ and $D = 2$. However, as Figs. 7 and 11 show, the computational complexity (the number of multiplications) of ASA is about 90% less than that of MLSD-4 with $D = 1$ at SNR $= 25$ dB and 93% less than that of MLSD-4 with $D = 2$ at SNR $= 20$ dB.

As seen in Figs. 6 and 10, the performances of ASA-DF2/1 (both $D = 1$ and $D = 2$) are between that of MLSD-2 and MLSD-3 at high SNR and less than the performance of MLSD-2 at low SNR. Moreover, the performances of ASA-DF2/2, which have not been shown in these figures, are worse than that of MLSD-2 for all SNR. This means that

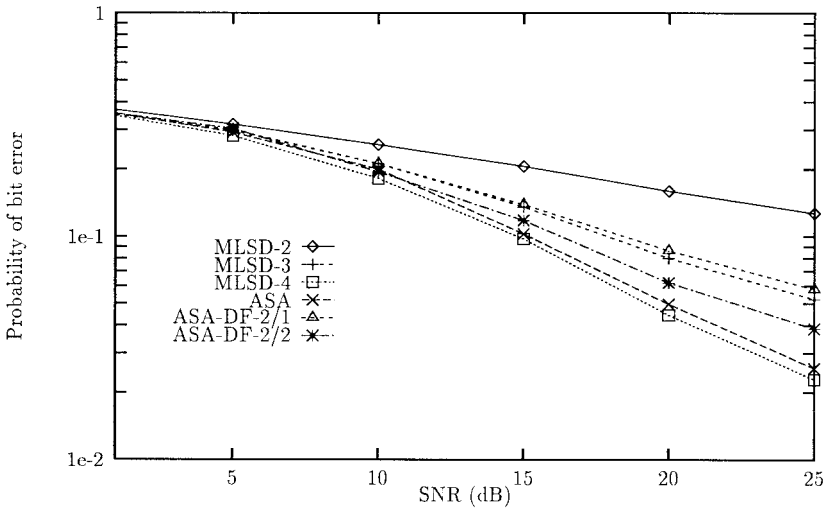


Fig. 8. Bit-error rate performance of selective fading with fivepaths for MLSD-2, MLSD-3, MLSD-4, ASA, ASA-DF-1/2, and ASA-DF-2/2 with $D = 1$ for $f_dT = 0.1$.

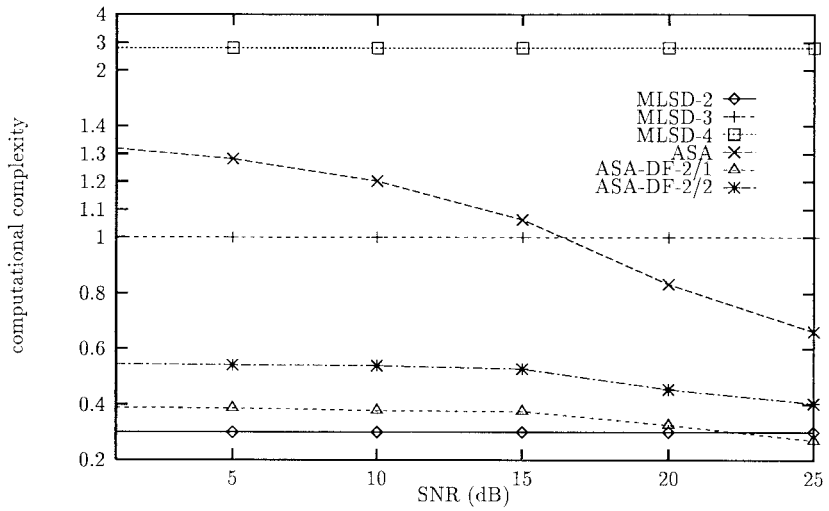


Fig. 9. Computational complexity of selective fading with five paths for MLSD-2, MLSD-3, MLSD-4, ASA, ASA-DF-1/2, and ASA-DF-2/2 with $D = 1$ for $f_dT = 0.1$. The computational complexity of all systems is normalized to that of MLSD-3 with $D = 1$.

by increasing the number of decision feedback symbols, the effects of error propagation degrade the performance when the channel goes to fade. Also, in flat fading, since there is no ISI and the correlation among symbols is only due to the fading, the performance is very sensitive to the number of decision feedback symbols. Figs. 7 and 11 show that the computational complexities of ASA-DF2/1 for flat fading are about 80% less than that of MLSD-3 and 93% less than that of MLSD-4 at SNR $= 25$ dB for $D = 1$ and at SNR $= 17.5$ dB for $D = 2$. Therefore, using the ASA-DF instead of the ASA in flat fading cannot be justified unless it is a high SNR situation and when the computational complexity issue is more important than the performance issue.

The performances of MLSD-2, MLSD-3, MLSD-4, ASA, ASA-DF2/1, and ASA-DF2/2 for selective fading with $D = 1$ and $D = 2$ are shown in Figs. 8 and 12, respectively. As results show in selective fading, the performances of ASA (both $D = 1$ and $D = 2$) are also close to that of MLSD-4, but as shown in Figs. 9 and 13 the computational complexities of ASA with $D = 1$ and $D = 2$ are about 77% and 85% less than that of MLSD-4 at SNR $= 25$ dB with $D = 1$ and $D = 2$, respectively.

In contrast to the flat fading channel, due to ISI in selective fading channels, the performance of ASA-DF2/2 in selective fading is better than ASA-DF2/1 (see Figs. 8 and 12) and it is between MLSD-3 and MLSD-4. As shown in Figs. 9 and 13, the computational complexities of ASA-DF2/1 are about 73% and 80% less than MLSD-3 for $D = 1$ and $D = 2$, respectively. Also, the computational complexities of ASA-DF2/2 are about 86% and 90% less than that of MLSD-4 for $D = 1$ and $D = 2$, respectively. Hence, using the ASA-DF instead of the ASA in selective fading is a tradeoff between computational complexity and performance.

All the results in the ASA algorithm show that the computational complexity is a function of SNR and increasing SNR decreases computational complexity. Also, in all systems, the difference between $D = 1$ and $D = 2$ is very large, hence, using diversity in fading channels is a very effective way to

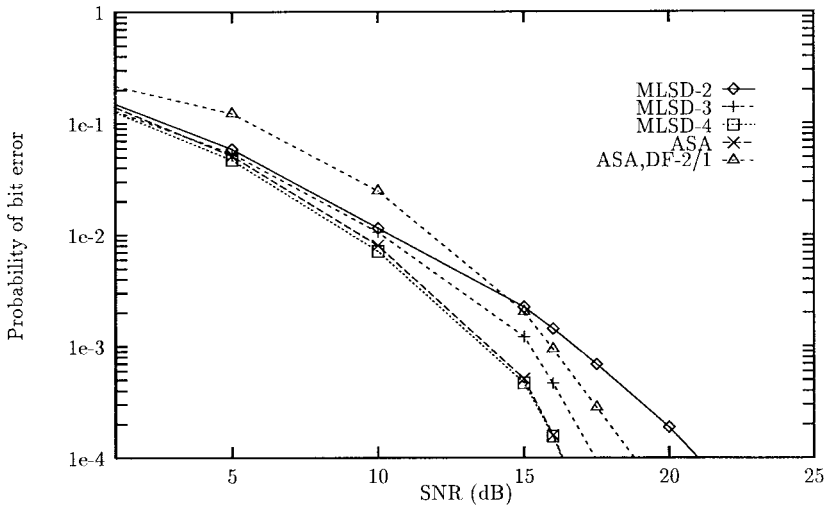


Fig. 10. Bit-error rate performance of flat fading for MLSD-2, MLSD-3, MLSD-4, ASA, and ASA-DF-1/2 with $D = 2$ and for $f_dT = 0.1$.

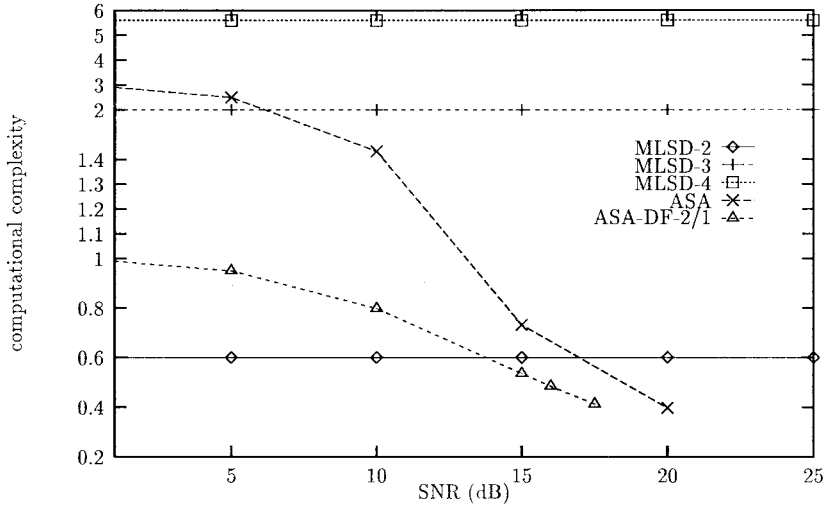


Fig. 11. Computational complexity of flat fading for MLSD-2, MLSD-3, MLSD-4, ASA, and ASA-DF-1/2 with $D = 2$ for $f_dT = 0.1$. The computational complexity of all systems is normalized to that of MLSD-3 with $D = 1$.

improve performance as well as to decrease the computational complexity of the ASA algorithm.

VII. Conclusion

We have described a new adaptive algorithm called the ASA algorithm which greatly reduces the computational complexity of the MLSD receiver based on measuring the short-term received signal power. In this algorithm, with adaptive state partitioning and adaptive threshold based on the Chernoff distance between the pdf's of correct and incorrect branch metrics, only a few states of the trellis diagram are chosen as the more likely correct states. Although the ASA algorithm is suboptimal, simulation results show that its performance is very close to that of the MLSD. The development of the algorithm shows clearly the relation between channel SNR, complexity and performance.

The tradeoff between computational complexity and the performance can be also controlled by the number of multiplications in each branch metric by using decision feedback symbols in the ASA (ASA-DF). Due to the absence of ISI in flat fading, the effects of error propagation limit the improvement in performance that can be achieved by increasing the number of decision feedback symbols. However, in selective fading, ASA-DF shows a good tradeoff between performance and computational complexity.

Computer simulations show that using diversity reception effectively improves the performance of MLSD, ASA and ASA-DF and decreases the computational complexity of ASA and ASA-DF as well.

Appendix A

The $R_{Y^{J-1}}$ is a symmetric positive definite matrix and similar to R_{Y^J}, it can be factorized as $R_{Y^{J-1}} = L_{J-1}(\tilde{\mathbf{a}}_k)D_{J-1}(\tilde{\mathbf{a}}_k)L^{\mathcal{H}}_{J-1}(\tilde{\mathbf{a}}_k)$, where $L_{J-1}(\tilde{\mathbf{a}}_k)$ is a unit upper triangular matrix and $D_{J-1}(\tilde{\mathbf{a}}_k)$ is a diagonal matrix, $D_{J-1}(\tilde{\mathbf{a}}_k) = \text{diag}(d_1(\tilde{\mathbf{a}}_k), d_2(\tilde{\mathbf{a}}_k), \cdots, d_{J-1}(\tilde{\mathbf{a}}_k))$. The inverse of $R_{Y^{J-1}}$ is $R^{-1}_{Y^{J-1}} = \Gamma^{\mathcal{H}}_{J-1}(\tilde{\mathbf{a}}_k)D^{-1}_{J-1}(\tilde{\mathbf{a}}_k)\Gamma_{J-1}(\tilde{\mathbf{a}}_k)$, where $\Gamma_{J-1}(\tilde{\mathbf{a}}_k) = L^{-1}_{J-1}(\tilde{\mathbf{a}}_k)$. From (19), one can easily show that all ijth elements of $\Gamma_{J-1}(\tilde{\mathbf{a}}_k)$ and $D^{-1}_{J-1}(\tilde{\mathbf{a}}_k)$ are

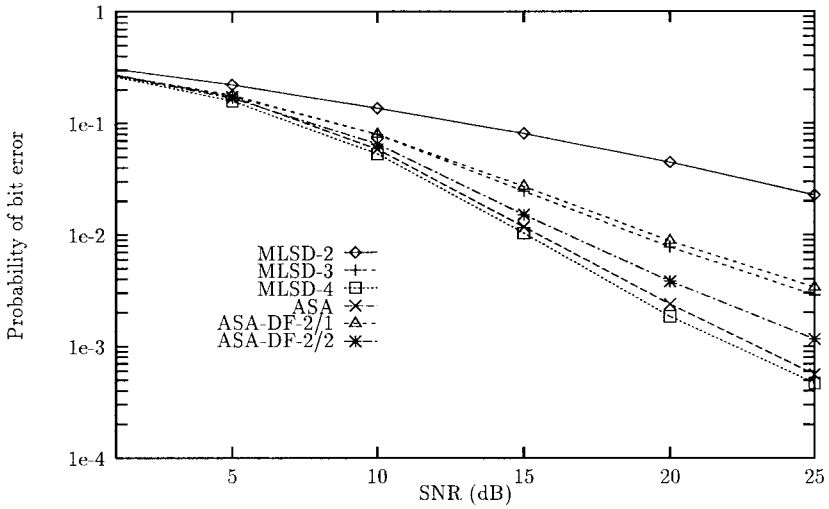


Fig. 12. Bit-error rate performance of selective fading with fivepaths for MLSD-2, MLSD-3, MLSD-4, ASA, ASA-DF-1/2, and ASA-DF-2/2 with $D = 2$ for $f_dT = 0.1$.

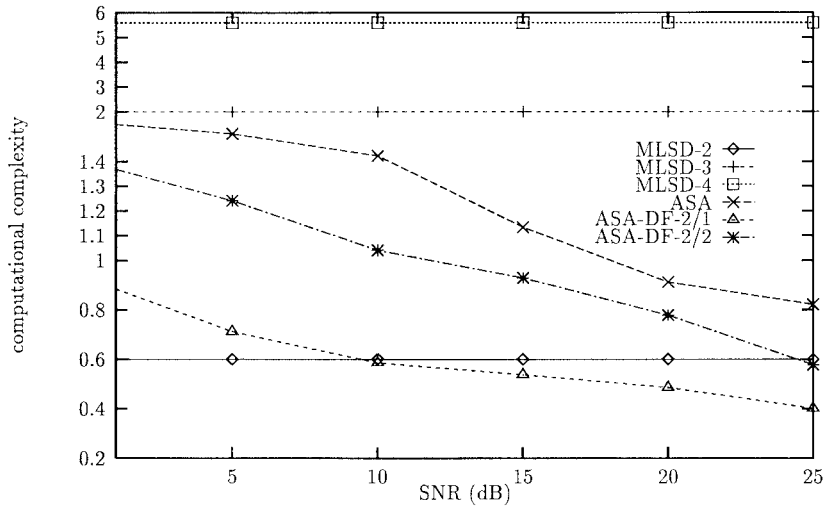


Fig. 13. Computational complexity of selective fading with five paths for MLSD-2, MLSD-3, MLSD-4, ASA, ASA-DF-1/2, and ASA-DF-2/2 with $D = 2$ for $f_dT = 0.1$. The computational complexity of all systems is normalized to that of MLSD-3 with $D = 1$.

equal to the all $(i-1)(j-1)$th elements of $\Gamma_J(\tilde{\mathbf{a}}_k)$ and $D_J^{-1}(\tilde{\mathbf{a}}_k)$ for $1 \le i, j \le J-1$, respectively. Then

$$\begin{bmatrix} 0 & \vdots & 0_{1\times(J-1)} \\ \cdots\cdots & \cdots & \cdots\cdots \\ 0_{(J-1)\times 1} & \vdots & R_{Y^{J-1}}^{-1} \end{bmatrix} = \Gamma_J^{\mathcal{H}}(\tilde{\mathbf{a}}_k) \begin{bmatrix} 0 & \vdots & 0_{1\times(J-1)} \\ \cdots\cdots & \cdots & \cdots\cdots \\ 0_{(J-1)\times 1} & \vdots & D_{Y^{J-1}}^{-1}(\tilde{\mathbf{a}}_k) \end{bmatrix} \Gamma_J(\tilde{\mathbf{a}}_k) \quad \text{(A-1)}$$

and $Q(\tilde{\mathbf{a}}_k)$ becomes

$$Q(\tilde{\mathbf{a}}_k) = \Gamma_J^{\mathcal{H}}(\tilde{\mathbf{a}}_k) \begin{bmatrix} d_0^{-1}(\tilde{\mathbf{a}}_k) & & & \\ & 0 & & 0 \\ 0 & & \ddots & \\ & & & 0 \end{bmatrix} \Gamma_J(\tilde{\mathbf{a}}_k)$$

$$= \begin{bmatrix} 1 \\ \gamma_{1,0}^*(\tilde{\mathbf{a}}_k) \\ \vdots \\ \gamma_{J-2,0}^*(\tilde{\mathbf{a}}_k) \\ \gamma_{J-1,0}^*(\tilde{\mathbf{a}}_k) \end{bmatrix} d_0^{-1}(\tilde{\mathbf{a}}_k) \times [1, \gamma_{1,0}(\tilde{\mathbf{a}}_k), \cdots, \gamma_{J-2,0}(\tilde{\mathbf{a}}_k), \gamma_{J-1,0}(\tilde{\mathbf{a}}_k)] \quad \text{(A-2)}$$

where $\gamma_{i,j}(\tilde{\mathbf{a}}_k)$ is the ijth element of $\Gamma_J(\tilde{\mathbf{a}}_k)$. Meanwhile the determinants R_{Y^J} and $R_{Y^{J-1}}$ are

$$\det(R_{Y^J}) = \prod_{j=0}^{J-1} d_j(\tilde{\mathbf{a}}_k)$$

$$\det(R_{Y^{J-1}}) = \prod_{j=1}^{J-1} d_j(\tilde{\mathbf{a}}_k).$$

Therefore, by substituting $Q(\tilde{\mathbf{a}}_k)$, $\det(R_{Y^J})$ and $\det(R_{Y^{J-1}})$ in (18) the branch metrics $B(\tilde{\mathbf{a}}_k)$ become (21)

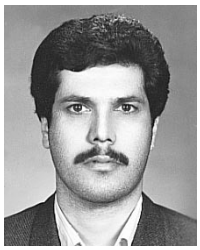
$$B(\tilde{\mathbf{a}}_k) = \ln(d_0(\tilde{\mathbf{a}}_k)) + \frac{\left|\sum_{j=0}^{J-1} y_{k-j}\gamma_{j,0}(\tilde{\mathbf{a}}_k)\right|^2}{d_0(\tilde{\mathbf{a}}_k)}. \qquad \text{(A-3)}$$

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their helpful comments and suggestions.

REFERENCES

[1] J. H. Lodge and M. L. Moher, "Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 38, pp. 787–794, June 1990.
[2] Q. Dai and E. Shwedyk, "Detection of bandlimited signals over frequency selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 941–950, Feb. 1994.
[3] X. Yu and S. Pasupathy, "Innovations-based MLSE for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 1534–1544, Feb./Mar./Apr. 1995.
[4] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.
[5] S. U. H. Qureshi and E. E. Newhall, "An adaptive receiver for data transmission over time-dispersive channels," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 448–457, July 1973.
[6] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Syst. Tech. J.*, vol. 9, pp. 1541–1562, Nov. 1973.
[7] C. T. Beare, "The choice of the desired impulse response in combined linear-Viterbi algorithm equalizers," *IEEE Trans. Commun.*, vol. COM-26, pp. 1301–1307, Aug. 1978.
[8] W. U. Lee and F. S. Hill, Jr., "A maximum-likelihood sequence estimator with decision-feedback equalization," *IEEE Trans. Commun.*, vol. COM-25, pp. 971–979, Sept. 1977.
[9] K. Wesolowski, "An efficient DFE and ML suboptimum receiver for data transmission over dispersive channels using two-dimensional signal constellations," *IEEE Trans. Commun.*, vol. COM-35, pp. 336–339, Mar. 1987.
[10] F. L. Vermeulen and M. E. Hellman, "Reduced state Viterbi decoders for channels with intersymbol interference," in *Proc. Int. Conf. Communications*, Minneapolis, MN, June 1974, pp. 37B-1–37B-4.
[11] G. J. Foschini, "A reduced state variant of maximum likelihood sequence detection attaining optimum performance for high signal-to-noise ratios," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 605–609, Sept. 1977.
[12] J. B. Anderson, "Limited search trellis decoding of convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 944–955, Sept. 1989.
[13] S. J. Simmons, "Breadth-first trellis decoding with adaptive effort," *IEEE Trans. Commun.*, vol. 38, pp. 3–12, Jan. 1990.
[14] S. J. Simmons and P. Senyshyn, "Reduced-search trellis decoding of coded modulation over ISI channels," in *Proc. IEEE Global Telecommunications Conf.*, San Diego, CA, Nov. 1990, pp. 401.7.1–401.7.4.
[15] J. B. Anderson and S. Mohan, "Sequential coding algorithm: A survey and cost analysis," *IEEE Trans. Commun.*, vol. COM-32, pp. 169–176, Feb. 1984.
[16] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, pp. 13–20, Jan. 1988.
[17] ———, "Reduced-state sequence estimation for coded modulation on intersymbol interference channels," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 989–995, Aug. 1989.
[18] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.
[19] P. R. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. Commun.*, vol. 37, pp. 669–676, July 1989.
[20] S. Olcer, "Reduced-state sequence detection of multilevel partial-response signals," *IEEE Trans. Commun.*, vol. 40, pp. 3–6, Jan. 1992.
[21] B. E. Spinnler and J. B. Huber, "Design of hyper states for reduced-state sequence estimation," in *Proc. IEEE Global Telecommunications Conf.*, San Diego, CA, Nov. 1995, pp. 1–6.
[22] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
[23] S. Stein, "Fading channel issues in system engineering," *IEEE J. Select. Areas Commun.*, vol. 40, pp. 68–89, Feb. 1987.
[24] E. A. Lee and D. G. Messerschmitt, *Digital Communications*, 2nd ed. Boston, MA: Kluwer, 1994.
[25] A. Papoulis, "Markoff and wide-sense Markoff sequences," in *Proc. IEEE*, p. 1661, Oct. 1965.
[26] ———, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1991.
[27] J. Makhoul, "A class of all-zero lattice digital filters: Properties and applications," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 304–314, Aug. 1978.
[28] T. L. M. Longo and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inform. Theory*, vol. IT-36, pp. 241–255, Mar. 1990.



Hossein Zamiri-Jafarian was born in Mashhad, Iran, in 1960. He received the B.Sc. and M.S. degrees from Isfahan University of Technology, Isfahan, Iran, in 1985 and 1987, respectively, and the Ph.D. degree from the University of Toronto, Toronto, Ont., Canada, in 1998, all in electrical engineering.

He is currently working as a Research Associate at the Department of Electrical and Computer Engineering, University of Toronto. He joined Ferdowsi University, Mashhad, as a Lecturer in 1987. His current research interests include statistical signal processing, data communications, and multirate signal processing.

Subbarayan Pasupathy (M'73–SM'83–F'91) was born in Chennai (Madras), Tamilnadu, India, on September 21, 1940. He received the B.E. degree in telecommunications from the University of Madras, Madras, India, in 1963, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, in 1966, and the M.Phil. and Ph.D. degrees in engineering and applied science from Yale University, New Haven, CT, in 1970 and 1972, respectively.

He joined the faculty of the University of Toronto, Toronto, Ont., Canada, in 1973 and became a Professor of Electrical Engineering in 1983. He has served as the Chairman of the Communications Group and as the Associate Chairman of the Department of Electrical Engineering, University of Toronto. He was an Associate Editor for the *Canadian Electrical Engineering Journal* from 1980 to 1983. His research interests are in the areas of communication theory, digital communications, and statistical signal processing.

Dr. Pasupathy was an Editor for Data Communications and Modulation for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1982 to 1989. He also served as a Technical Associate Editor for the IEEE COMMUNICATIONS MAGAZINE from 1979 to 1982. Since 1984, he has been writing a regular column entitled "Light Traffic" for the IEEE COMMUNICATIONS MAGAZINE. He is a registered Professional Engineer in the Province of Ontario, Canada.