Complexity Reduction of the MLSD/MLSDE Receiver Using the Adaptive State Allocation Algorithm

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

***Abstract*—The idea of adaptive state allocation (ASA) algorithm is used in this paper to substantially reduce the computational complexity of the maximum-likelihood sequence detection and estimation (MLSD/MLSDE) receiver without a significant degradation in its performance. In the ASA algorithm, the total number of states assigned to the trellis and the number of states selected from the entire set are changed adaptively based on the short-term power of the channel impulse response (CIR) or its estimate.**

The ASA algorithm is a combination of two methods: adaptive threshold (AT) and adaptive state partitioning (AP). In the AT method, a threshold value is formulated based on the probability of removing the correct state in the trellis diagram. At each time, only the paths whose costs are less than the minimum cost (corresponding to the best survivor path) plus the threshold value are retained and are extended to the next trellis stage. The AT method significantly reduces the computational complexity of the regular MLSDE mostly at high signal-to-noise ratio (SNR) with a negligible loss in performance. Simulation results for fading channels show that the AT method typically selects one trellis state (the minimum possible number of states) at high SNRs. In the AP method, the branch metrics are fused and diffused adaptively by using the Kullback–Leibler (KL) distance metric invoked for quantifying the differences between the probability density functions of the correct and incorrect branch metrics in the trellis. The adaptation is done such that the channel coefficients with short-term power less than a threshold are assumed to be zero in computing the branch metrics. The AP method decreases the computational complexity of the regular MLSDE at low SNRs.

***Index Terms*—Adaptive detection and estimation, complexity reduction, complexity theory, fading channels, maximum-likelihood detection, multipath channels, sequence detection theory.**

I. Introduction

MAXIMUM likelihood sequence detection and estimation (MLSD/MLSDE) is an optimal receiver technique designed for intersymbol interference-contaminated multipath fading channels, which minimizes the sequence detection error rate [1], [2]. However, the major problem associated with such a nonlinear receiver is its computational complexity, which increases exponentially with the channel's memory. A considerable amount of research has been conducted in order to find less complex algorithms or structures exhibiting near-optimal performance [3]–[8].

There are three distinct approaches for reducing the complexity of the MLSD/MLSDE receiver employing the Viterbi (dynamic programming) algorithm. The first class of methods decreases the overall effective memory length of the channel/equalizer, in order to reduce the number of states in the trellis diagram and, hence, the number of required computations for each branch metric [3], [4]. Another approach is to select only a subset of the total number of trellis states, as proposed in the context of the reduced-state sequence estimation (RSSE) method [5], the M-algorithm [6] and the T-algorithm [7]. According to the RSSE method, a subset of the total number of states is selected such that the number of states in the subset and the states allocated to it are fixed. By contrast, in the context of the M-algorithm, only the number of selected states is fixed. In the T-algorithm, neither the number of states in the selected subset nor the states allocated to the selected subset is fixed. Based on this approach, other algorithms such as the modified RSSE [9], the bidirectional tree searching algorithm [10], the trellis search algorithm [11] and the reduced state decision feedback estimation algorithm [12] have been proposed. In [11], motivated by the M-algorithm, the computational complexity of a blind joint channel estimation and data detection receiver has been reduced and in [12] the T-algorithm is used for reducing the computational complexity of a suboptimal symbol-by-symbol detection scheme.

According to the third approach, which is referred to as the adaptive state allocation (ASA) algorithm [8], the total number of states in the trellis, the number of states in the selected subset and also the states allocated to the selected subset are changed adaptively. Therefore, the number of branch metrics and the number of computations required for each branch metric are time-variant in the ASA algorithm. This approach is more suitable for a time-variant environment, such as a multipath fading channel which disperses the transmitted signal both in the time and frequency domains. Since the short-term power of the received signal is time-variant in multipath fading channels, the ASA algorithm succeeds in using the available computational budget in an efficient manner. In order to keep the receiver performance close to optimum, the ASA algorithm allocates more computational budget, when the short-term power of the channel impulse response (CIR) is low or the channel is in a deep fade, and conserves some of its computational budget when the channel quality is high.

The ASA algorithm in [8] is based on the statistical parameters of a random CIR and does not consider the case of knowing

Manuscript received September 28, 1999; revised July 19, 2000 and January 7, 2001; accepted January 7, 2001. The editor coordinating the review of this paper and approving it for publication is L. Hanzo. This work was presented in part at the IEEE International Conference on Communications (ICC), Vancouver, June 1999.

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

Publisher Item Identifier S 1536-1276(02)00189-7.

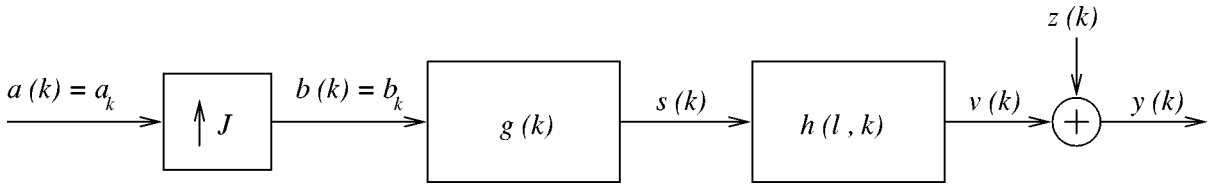


Fig. 1. The discrete model of the transmitter and the fading channel, where $\uparrow J$ represents up-sampling by a factor J.

the CIR or its estimate. In this paper, the ASA algorithm of [8] is extended to the situation in which the CIR or its estimate is known, such as the situation in a MLSD/MLSDE receiver [2]. The MLSDE receiver, which uses a joint CIR estimation and data detection method along with per-survivor processing or per-branch processing [2], achieves good performance due to its higher number of computations; hence, reducing the computational complexity of this type of receivers is more crucial in a practical sense.

The extended ASA algorithm selects only a few trellis states based on an adaptive threshold (AT) method [13] and fuses branch metrics by using an adaptive partitioning (AP) method for reducing the computational complexity. At each trellis stage, the AT method selects the states having a cost less than a threshold plus the minimum cost, where the latter is the sum of the branch metrics in the best survivor path. The threshold value is formulated based on the probability of removing the correct states (or the probability of suspending the correct path) in the trellis. This probability is chosen to be a small value. The idea of selecting a few states based on a threshold value according to the T-algorithm has been proposed in [7] without giving a formula for calculating such a threshold. The AP method is based on the Kullback–Leibler (KL) distance metric invoked for quantifying the difference between the probability density functions of the correct and the incorrect branch metrics in the trellis diagram. In the AP method, due to the time-variant properties of the channel, the CIR taps or channel coefficients having a short-term power lower than the threshold are assumed to be zero in calculating the branch metrics.

Although the extended ASA algorithm intends to reduce the computational complexity of the detection part of the MLSDE receiver, the complexity of the estimation part is decreased as well due to the reduction in the number of branches, for which estimation of the CIR is necessary.

This paper is organized as follows. The system model and a brief derivation of the MLSD/MLSDE receiver structure is presented in Section II. The extension of the ASA algorithm for the receiver which knows the CIR or its estimate is developed in Section III. Some notes about the implementation of the ASA are provided in Section IV. Section V contains some simulation results and comparisons for frequency flat and selective fading channels. Finally, some conclusions are presented in Section VI.

II. MLSD/MLSDE Receiver

The discrete model of a system communicating over a fading channel is shown in Fig. 1, where a_k is the transmitted symbol, which is taken from a q-ary constellation, while b_k is the up-sampled version of a_k by a factor J. Furthermore, $g(k)$ is the impulse response of the transmitter filter, $s(k)$ is the transmitted signal, and $h(l,k) = h_l(k)$ is the CIR, while $z(k)$ is the additive white Gaussian noise having a variance of N_0 and $y(k)$ is the received signal.

Detecting $\mathbf{a} = (a_0, \ldots, a_{\mathrm{I}-1})$ based on the MLSD criterion is given by

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}\in\mathcal{A}}\{p(\mathbf{y}|\mathbf{a})\} = \arg\max_{\mathbf{a}\in\mathcal{A}}\{\log p(\mathbf{y}|\mathbf{a})\} \tag{1}$$

where $\hat{\mathbf{a}}$ is the detected estimate of $\mathbf{a}$, $\mathcal{A}$ is the set of all possible transmitted symbol sequences and $\mathbf{y} = [y_0, \ldots, y_{K-1}]^{\mathcal{T}}$ is the sequence of the received signal, while $X^{\mathcal{T}}$ denotes the transpose of X and K is the total number of received samples. It can be shown that after some manipulations (1) becomes [14]

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}\in\mathcal{A}}\left\{\sum_k \beta(k)\right\} \tag{2}$$

where $\beta(k)$ is the branch metric at time k. When the CIR is known $\beta(k)$ is given by

$$\beta(k) = |y_k - \boldsymbol{s}(k)\boldsymbol{h}(k)|^2 \tag{3}$$

where $\boldsymbol{s}(k) = [s(k), \ldots, s(k-L)]$, L is the memory length of the channel, and $\boldsymbol{h}(k) = [h_0(k), \ldots, h_L(k)]^{\mathcal{T}}$. However, when $\boldsymbol{h}(k)$ is not available, the estimate of $\boldsymbol{h}(k)$, $\hat{\boldsymbol{h}}(k)$, is used in (3) instead of $\boldsymbol{h}(k)$ and the receiver is termed as the MLSDE due to the joint channel estimation and data detection based on the maximum-likelihood criterion [2].

III. The ASA Algorithm

The main goal of the ASA algorithm is to reduce the computational complexity of the regular MLSD/MLSDE, such that its performance remains close to that of the regular MLSD/MLSDE. The performance of the regular MLSD/MLSDE implemented using the Viterbi algorithm depends on $d_{\min}$, the minimum distance error in the trellis diagram, as defined in [15, p. 418]. In a time-variant environment, $d_{\min}$ is also time-variant due to changes in the channel coefficients or the taps of the CIR. The performance of the regular MLSD/MLSDE is dominated by $d_{\min}$. Therefore, in order to keep the performance of the ASA algorithm close to that of the regular MLSD/MLSDE receiver, the computational complexity should be reduced such that the possibility of selecting the wrong path instead of the correct path becomes negligible due to changes made by the ASA algorithm in the trellis structure. Based on this goal, the AT and AP methods are developed in Sections III-A and B, when the CIR or its estimate is available at the receiver.

A. Adaptive Threshold (AT) Method

One method of reducing the computational complexity of the MLSD/MLSDE is to remove some states whose costs[1] are larger than a threshold value. This idea, which is referred as the T-algorithm, was proposed in [7] for reducing the decoding complexity of convolutional codes. However, in [7] no formula was suggested for selecting the threshold value. In [7], the relation between the performance and decoding complexity was studied based on computer simulations by selecting different threshold values. We use the concept proposed in [7] in order to reduce the computational complexity of the MLSD/MLSDE receiver designed for detecting the transmitted signal over the multipath fading environment such that the performance of the reduced complexity receiver remains close to that of the regular one. At each time the states whose costs are larger than a threshold plus a minimum cost (the cost of the best survivor path) are removed from the trellis diagram. The strategy of choosing the threshold value is considered in the remainder of this section.

Assuming that the CIR is known to the receiver, the branch metric between the ith state and the jth state at time k is

$$\beta_{ij}(k) = |y(k) - \boldsymbol{s}_{ij}(k)\boldsymbol{h}(k)|^2 \tag{4}$$

where $\boldsymbol{s}_{ij}(k)$ is a row vector of the transmitter filter output samples corresponding to transmission symbols relating the ith state to the jth state. Let us assume that the ith state is the correct state at which the correct path and incorrect paths begin to diverge. We define $\beta_i^c(k)$ as the correct branch metric and $\beta_i^u(k)$ as the minimum incorrect branch metric, which diverges from the ith state at time k (Fig. 2). Since both correct and incorrect branches are diverging from the same state, the probability of removing the correct state (or suspending the correct path) is[2]

$$p_c^r = p\left(\beta_i^c(k) > \beta_i^u(k) + T_{h_i}(k)\right) \tag{5}$$

where $T_{h_i}(k)$, the threshold, is a nonnegative value. The more likely correct states are selected based on $T_{h_i}(k)$ when the minimum distance error event has only two branches (Fig. 2). From (4), the parameters $\beta_i^c(k)$ and $\beta_i^u(k)$ are

$$\begin{aligned}\beta_i^c(k) =& \beta_{ij}(k) = |z(k)|^2 \qquad (6)\\ \beta_i^u(k) =& \beta_{il}(k) = |z(k) + \Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2\\ =& |z(k)|^2 + |\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2\\ &+ 2(z_{\mathrm{R}}(k)\mathrm{Re}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)]\\ &+ z_{\mathrm{I}}(k)\mathrm{Im}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)]) \qquad (7)\end{aligned}$$

where $z_{\mathrm{R}}(k)$ and $z_{\mathrm{I}}(k)$ are the real and imaginary parts of $z(k)$, respectively, while $\Delta\boldsymbol{s}_i(k) = \boldsymbol{s}_i^c(k) - \boldsymbol{s}_i^u(k)$ such that $\boldsymbol{s}_i^c(k) = \boldsymbol{s}_{ij}(k)$ and $\boldsymbol{s}_i^u(k) = \boldsymbol{s}_{il}(k)$ are the transmitter filter outputs corresponding to the correct and the incorrect branch metrics, and

[1]The cost of each state is the sum of the branch metrics in its survivor path.

[2]We assume that the correct state is not the minimum cost state in (5). Therefore, computing $T_{h_i}(k)$ from (5) decreases the probability of removing the correct state.

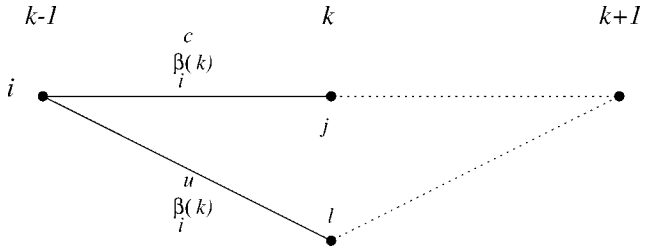


Fig. 2. The minimum distance error event with two branches originating from the ith state.

$\mathrm{Re}[X]$ and $\mathrm{Im}[X]$ denote the real and imaginary parts of X, respectively. By replacing (6) and (7) in (5), we have

$$\begin{aligned}p_c^r =& p\Big(2\Big(z_{\mathrm{R}}(k)\mathrm{Re}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)]\\ &+ z_{\mathrm{I}}(k)\mathrm{Im}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)]\Big)\\ &< -\left(|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2 + T_{h_i}(k)\right)\Big)\end{aligned} \tag{8}$$

where $z(k)$ is a circularly symmetric zero-mean white complex Gaussian random process with autocorrelation $R_z(\tau) = N_0\delta(\tau)$, while $z_{\mathrm{R}}(k)$ and $z_{\mathrm{I}}(k)$ are independent zero-mean Gaussian random processes, each with a variance of $N_0/2$. Therefore, $2(z_{\mathrm{R}}(k)\mathrm{Re}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)] + z_{\mathrm{I}}(k)\mathrm{Im}[\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)])$ becomes zero-mean Gaussian noise with variance $2N_0|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2$ and p_c^r is given by

$$\begin{aligned}p_c^r =& \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\left(|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2 + T_{h_i}(k)\right)^2}{4N_0|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2}}\right)\\ =& \frac{1}{2}\mathrm{erfc}(C_1)\end{aligned} \tag{9}$$

where $\mathrm{erfc}(.)$ denotes the complementary error function and $C_1 \doteq \sqrt{(|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2 + T_{h_i}(k))^2/(4N_0|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2)}$. By selecting a value for p_c^r, which should be very much smaller than one, it is easy to find C_1, the argument of the complementary error function. Thus, the nonnegative threshold value at time k, $T_{h_i}(k)$, becomes

$$\begin{aligned}T_{h_i}(k) = \max\Big(&2C_1\sqrt{N_0|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2}\\ &-|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2, 0\Big).\end{aligned} \tag{10}$$

In developing (10), it was assumed that the CIR, $\boldsymbol{h}(k)$, was known and $\Delta\boldsymbol{s}_i(k)$ is a deterministic vector. However, only the estimate of the CIR is available in an adaptive MLSDE and the value of $\Delta\boldsymbol{s}_i(k)$ depends on the state and branch metric, which are assumed to be the correct ones. When the estimation is reliable, one can substitute $\hat{h}(k)$, the estimate of $h(k)$, instead of $h(k)$ in (10); then the expectation of $|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2$ can be approximated by

$$\begin{aligned}E\left[|\Delta\boldsymbol{s}_i(k)\boldsymbol{h}(k)|^2\right] \simeq& E\left[|\Delta\boldsymbol{s}_i(k)\hat{\boldsymbol{h}}(k)|^2\right]\\ =& \hat{\boldsymbol{h}}^{\mathcal{T}}(k)E\left[\Delta\boldsymbol{s}_i^{\mathcal{T}}(k)\Delta\boldsymbol{s}_i^*(k)\right]\hat{\boldsymbol{h}}^*(k)\end{aligned} \tag{11}$$

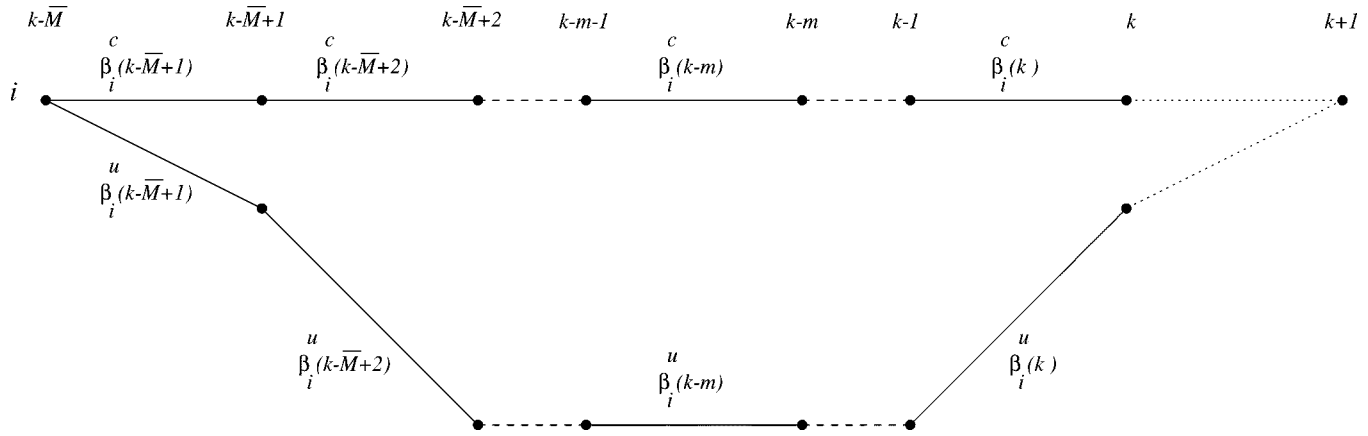


Fig. 3. The minimum distance error event originating from the ith state in the trellis diagram.

where X^* denotes the conjugate of X. $E\left[\Delta \boldsymbol{s}_i^T(k)\Delta \boldsymbol{s}_i^*(k)\right]$ can be calculated by assuming that all states have the same probability of being the correct state (before selecting some of them as the more likely correct states). Therefore, the average threshold for selecting the more likely correct states is approximated by

$$\bar{T}_h(k) \simeq \max\left(2C_1\sqrt{N_0 E\left[|\Delta \boldsymbol{s}_i(k)\hat{\boldsymbol{h}}(k)|^2\right]} - E\left[|\Delta \boldsymbol{s}_i(k)\hat{\boldsymbol{h}}(k)|^2\right], 0\right). \tag{12}$$

As mentioned before, the error probability of the MLSDE for channels exhibiting memory is dominated by the minimum distance error event in the trellis diagram. The $\bar{T}_h(k)$ value in (12) is obtained based on the distance of one branch between the correct path and the incorrect path. In other words, it is assumed that the minimum distance error event in the trellis diagram contains only two branches. When the length of the minimum distance error event contains $\bar{M}+1$ branches, the probability of removing the correct state at the mth branch of the minimum distance error event is

$$p(\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{m-1},\varepsilon_m) = p(\varepsilon_m|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{m-1})p(\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{m-1}) \quad m \le \bar{M} \tag{13}$$

where

$\varepsilon_m \doteq$ removing the correct path at the mth branch of the minimum distance error event;

$\bar{\varepsilon}_m \doteq$ retaining the correct path at the mth branch of the minimum distance error event.

It can be seen from Fig. 3 that when the sum of $\bar{m}$ branch metrics in the correct path is larger than the sum of $\bar{m}$ branch metrics in the incorrect path plus the threshold, the correct path will be removed. Therefore, $p(\varepsilon_{\bar{m}}|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1})$, the probability of removing the correct path at the $\bar{m}$th branch of the minimum distance error event given that it is retained up to the $(\bar{m}-1)$th branch, becomes

$$p(\varepsilon_{\bar{m}}|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1}) = p\left(\sum_{m=0}^{\bar{m}-1}\beta_i^c(k-m) > \sum_{m=0}^{\bar{m}-1}\beta_i^u(k-m) + T_{h_i}(k)\right) \quad \bar{m} \le \bar{M}. \tag{14}$$

Based on the definition of the correct and the incorrect branch metrics, it is easy to show that

$$\sum_{m=0}^{\bar{m}-1}\beta_i^c(k-m) = \sum_{m=0}^{\bar{m}-1}|z(k-m)|^2 \tag{15}$$

$$\sum_{m=0}^{\bar{m}-1}\beta_i^u(k-m) = \sum_{m=0}^{\bar{m}-1}|z(k-m)|^2 + |\Delta \boldsymbol{s}_i(k-m)\boldsymbol{h}(k-m)|^2 + 2(z_{\mathrm{R}}(k-m)\mathrm{Re}[\Delta \boldsymbol{s}_i(k-m)\boldsymbol{h}(k-m)] + z_{\mathrm{I}}(k-m)\mathrm{Im}[\Delta \boldsymbol{s}_i(k-m)\boldsymbol{h}(k-m)]). \tag{16}$$

At each time, the tentative threshold value should be computed for all $1 \le \bar{m} \le \bar{M}$ and then the maximum value is used as the threshold. The average threshold value, $\bar{T}_h(k)$, can be approximated at time k by following the procedure that has been invoked for $\bar{M} = 1$ [see (12)], yielding

$$\bar{T}_h(k) \simeq \max\left(\max_{1\le\bar{m}\le\bar{M}}\left\{2C_1\sqrt{N_0\sum_{m=0}^{\bar{m}-1}E\left[|\Delta \boldsymbol{s}_i(k-m)\hat{\boldsymbol{h}}(k-m)|^2\right]} - \sum_{m=0}^{\bar{m}-1}E\left[|\Delta \boldsymbol{s}_i(k-m)\hat{\boldsymbol{h}}(k-m)|^2\right]\right\}, 0\right) \tag{17}$$

where, similar to the case of $\bar{M} = 1$, choice of the constant C_1 is based on the probability of removing the correct states in the trellis diagram. Meanwhile, $p(\varepsilon_{\bar{m}}|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1}) \ge p(\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1},\varepsilon_{\bar{m}})$; therefore, since we obtain a larger value by using $p(\varepsilon_{\bar{m}}|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1})$, the value of $T_h(k)$ based on $p(\varepsilon_{\bar{m}}|\bar{\varepsilon}_1,\ldots,\bar{\varepsilon}_{\bar{m}-1})$ is an upper-bound. The dynamic range of $\bar{T}_h(k)$, which varies according to the channel conditions, is shown in Fig. 4 where x is defined as

$$x = \left(\sum_{m=0}^{\bar{m}-1}E\left[|\Delta \boldsymbol{s}_i(k-m)\hat{\boldsymbol{h}}(k-m)|^2\right]\right)^{1/2}. \tag{18}$$

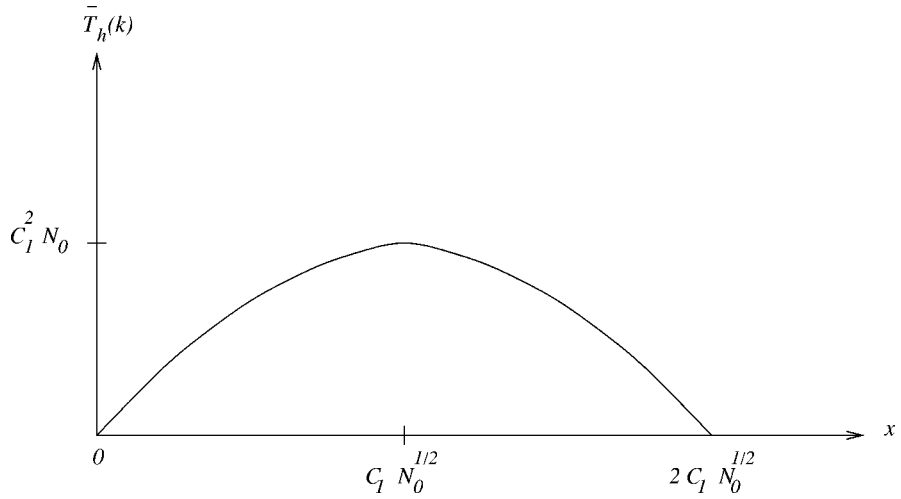


Fig. 4. $\bar{T}_h(k)$ as a function of the channel condition; x is defined in (18).

As Fig. 4 shows, the threshold value becomes very small for two different channel conditions, namely: when the channel quality is high or the received power is high and when the channel quality is low or the received power is low in the time duration of the minimum distance error event. In other words, when the channel stays in a good or bad condition for a duration longer than the time in which a minimum distance error event can happen, it is sufficient to retain only one state (minimum cost survivor path) in the trellis diagram and remove the other states.

The straightforward way of avoiding the calculation of $T_h(k)$ at each time is to select the maximum possible value of $T_h(k)$. From (17), it is easy to show that by considering the threshold as a function of x, its maximum value, $T_{h_{\max}}$, becomes

$$T_{h_{\max}} = C_1^2 N_0. \tag{19}$$

As the surprisingly simple formula (19) shows, the knowledge of the CIR and the transmitted signal is not necessary in order to compute the value of $T_{h_{\max}}$. It is only a function of C_1, which is found from the probability of removing the correct state, and N_0, the variance of the additive Gaussian noise. Moreover, $T_{h_{\max}}$ is given by (19) even if $|\Delta \boldsymbol{s}_i(k-m)\boldsymbol{h}(k-m)|^2$ is used in (17) instead of the expectation of its estimate; in this situation, x is defined as

$$x = \left(\sum_{m=0}^{\bar{m}-1} |\Delta \boldsymbol{s}_i(k-m)\boldsymbol{h}(k-m)|^2 \right)^{1/2}.$$

Therefore, the maximum value of the threshold in (19) is independent of which state is labeled as the ith state. In other words, regardless of the state from which the correct and incorrect paths diverge and without knowledge about the CIR or its reliable estimate, $T_{h_{\max}}$ in (19) can be considered as the maximum value for the threshold.[3]

Meanwhile, the AT method can be used in order to reduce the computational complexity of the MLSDE receiver for the standard-ISI-contaminated channel as well.

[3]It should be mentioned that for computing the exact value of the threshold [such as (10)] knowing the ith state and the CIR are necessary; only for calculating its maximum (19) such knowledge is not necessary.

B. Adaptive State Partitioning (AP) Method

The CIR at each tap, $h_l(k)$, is modeled as a Gaussian random process in the multipath fading channel. The number of states increases exponentially with "l" or the memory of the channel. One way of reducing the computational complexity of convolutional decoding and that of the MLSD receiver is to use the idea of state partitioning or to consider only a subset of all the possibilities for all the symbols [5]. Based on this idea, we propose the fusing and diffusing of branch metrics in the trellis diagram—a process which becomes explicit below—in association with the power of $h_l(k)$ for $l = 0, \ldots, L$. When the power of $h_l(k)$ for a specific "l" is less than a threshold, it is assumed to be zero. Therefore, some branch metrics, which are different from each other due to the nonzero values of the $h_l(k)$, become the same and they are fused. As an example, Fig. 5 shows a discrete multipath fading channel having an impulse response duration of $L = 3$. The corresponding trellis diagram of this channel for $q = 2$ (q is the number of symbols in the modulation constellation) is shown in Fig. 6. During the transition from time $k-1$ to time k, the sequence $(a_k, a_{k-1}, a_{k-2}, a_{k-3})$ corresponding to $\boldsymbol{s}(k) = [s(k), s(k-1), s(k-2), s(k-3)]$, and $h_l(k)$ for $l = 0, \ldots, 3$ contributes to the calculation of each branch metric.[4]

Assuming $h_l(k) = 0$ for each $0 \le l \le L$ reduces the number of branch metrics by a factor q and also decreases the number of multiplications in calculating each branch metric. For example, in Fig. 6, if it is assumed that $h_1(k) = 0$, the symbol a_{k-1} does not contribute to the branch metrics, therefore, the branch metrics (0,2), (1,3), (4,6), (5,7), (8,10), (9,11), (12,14), and (13,15) are fused and become equal[5] or if $h_3(k) = 0$, the symbol a_{k-3} has no role in calculating the branch metrics and the branch metrics (0,8), (1,9), (2,10), (3,11), (4,12), (5,13), (6,14), and (7,15) become equal. Therefore, all the possibilities of the symbol (or symbols), which correspond to $h_l(k) = 0$, are ignored in computing the branch metrics. Since the power of $h_l(k)$ is time-

[4]For the sake of simplicity, we have not considered the effects of the transmitter filter memory and have only considered the situation of one sample per symbol.

[5]Branch metrics (i, j) are said to be fused when ith branch metric $= j$th branch metric; see Fig. 6.

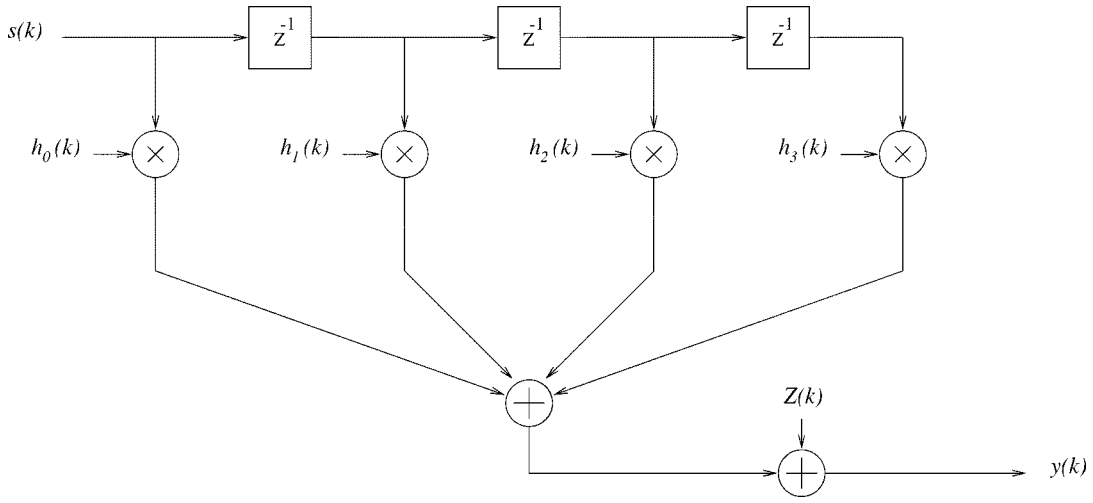


Fig. 5. The digital channel model with $L = 3$.

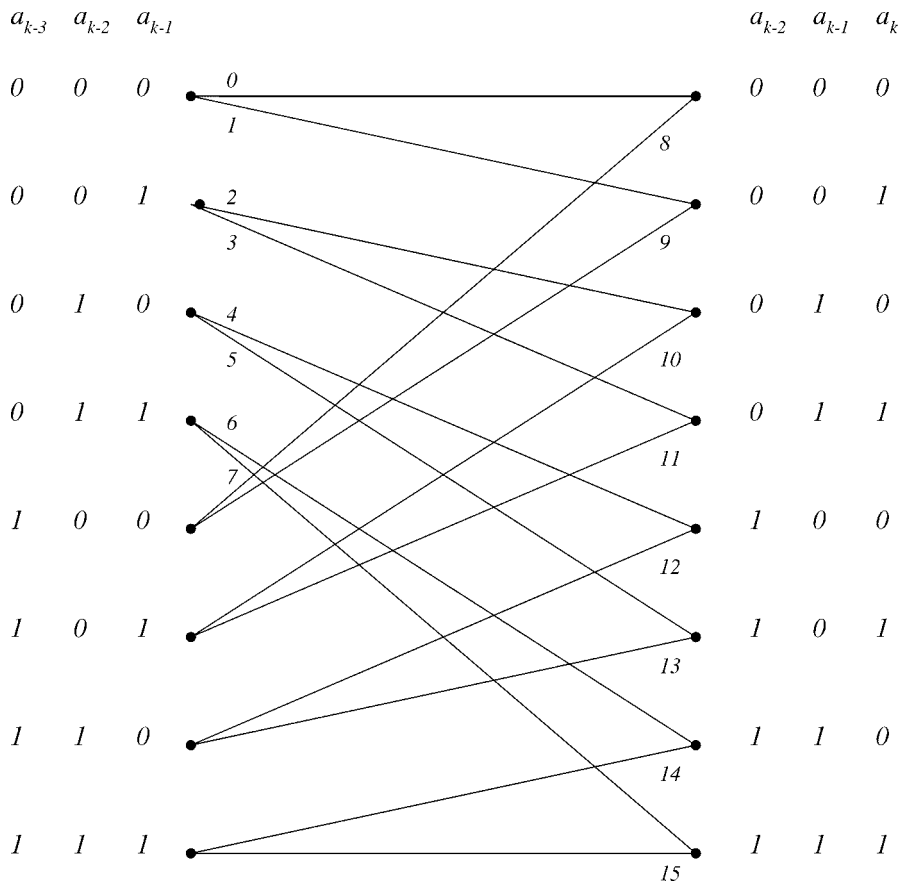


Fig. 6. The trellis diagram for the channel model shown in Fig. 5 with $q = 2$. The branches are indicated by the numbers $0, 1, \ldots, 15$.

variant, the partitioning of states is also time-variant based on the quality of the channel. Hence, we refer to this method as "adaptive state partitioning" or, briefly, AP.

The strategy of the AP method is to keep the error probabilities of the partitioned and the unpartitioned MLSDE receiver close to each other. The branch metric in the MLSDE receiver is considered as a random process which combines two stationary random processes $z(k)$ and $\boldsymbol{h}(k)$. Since $z(k)$ and $\boldsymbol{h}(k)$ are circularly symmetric zero-mean complex Gaussian random processes, which are independent of each other, the distribution of the branch metric of (4) is a two degree chi-square random process, whose density function for the correct and the incorrect branch metrics at each time k is given by

$$f_c(\beta) = \frac{1}{\sigma_c^2} e^{-(\beta(k)/\sigma_c^2)} \tag{20}$$

$$f_u(\beta) = \frac{1}{\sigma_u^2} e^{-(\beta(k)/\sigma_u^2)} \tag{21}$$

respectively. Similar to (6) and (7), $\beta^c(k) = |z(k)|^2$ and $\beta^u(k) = |z(k) + \Delta\boldsymbol{s}(k)\boldsymbol{h}(k)|^2$, therefore, σ_c^2 and σ_u^2 are

$$\sigma_c^2 = E[\beta^c(k)] = N_0 \tag{22}$$

$$\sigma_u^2 = E[\beta^u(k)] = N_0 + \Delta\boldsymbol{s}(k)E\left[\boldsymbol{h}(k)\boldsymbol{h}^{\mathcal{H}}(k)\right]\Delta\boldsymbol{s}^{\mathcal{H}}(k) \tag{23}$$

such that $\Delta\boldsymbol{s}(k) = [\Delta s_0(k), \ldots, \Delta s_L(k)]$ is the transmitted signal-difference between the correct and the incorrect branch metrics and $X^{\mathcal{H}}$ denotes the conjugate transpose of X. From the viewpoint of error event probability the worst case scenario is encountered, when the incorrect branch metric becomes the correct branch metric by fusion. In other words, the symbol (symbols) that are different for the correct and the incorrect branch metrics corresponding to the channel coefficient (coefficients), $h_l(k)$, becomes zero. We would like to find the relation between the distance of the correct and incorrect branch metrics and the power associated with the CIR taps. Due to the randomness of the branch metrics, we consider the distance between the probability density functions of the incorrect and the correct branch metrics. There are different definitions for the distance between two probability density functions [16]. Here, we use the KL distance[6] between $f_u(\beta)$ and $f_c(\beta)$ for assessing whether an assumed distribution is the true distribution (see [19, p. 18]).

$$d(\beta^u, \beta^c) = \int \log\left(\frac{f_u(\beta)}{f_c(\beta)}\right) f_u(\beta)d\beta = \frac{\Delta\sigma^2}{\sigma_c^2} - \log\left(1 + \frac{\Delta\sigma^2}{\sigma_c^2}\right) \tag{24}$$

where one can show, based on the mutual independence of elements $\boldsymbol{h}(k)$ and with the aid of (22) and (23), that

$$\Delta\sigma^2 = \sigma_u^2 - \sigma_c^2 = \sum_{l=0}^{L} |\Delta s_l(k)|^2 E\left[|h_l(k)|^2\right]. \tag{25}$$

In the minimum distance error event, only one symbol is different between the correct and the incorrect branch metrics. In other words, only one element of $\Delta\boldsymbol{s}(k)$ is nonzero in each branch. If we assume that the lth element of $\Delta\boldsymbol{s}(k)$, $\Delta s_l(k)$, is nonzero and use the average value of $|\Delta s_l(k)|^2$ and since $\Delta s_l(k)$ will be different for different symbols selected from a set of q possibilities, we have

$$\Delta\sigma_l^2 \doteq \overline{\Delta\sigma^2} = \overline{|\Delta s_l(k)|^2} E\left[|h_l(k)|^2\right]. \tag{26}$$

Therefore, the KL distance for lth error in the minimum distance error event becomes

$$d^l(\beta^u, \beta^c) = \frac{\Delta\sigma_l^2}{\sigma_c^2} - \log\left(1 + \frac{\Delta\sigma_l^2}{\sigma_c^2}\right) \tag{27}$$

where $\bar{d}(\beta^u, \beta^c)$, the average KL distance between the correct and the incorrect branch metrics in the minimum distance error event, is given by

$$\bar{d}(\beta^u, \beta^c) = \frac{1}{L+1}\sum_{l=0}^{L} d^l(\beta^u, \beta^c). \tag{28}$$

[6]Even though the KL measure does not satisfy all the properties of a proper distance, it has nevertheless been used widely in many similar problems (see, e.g., [17] and [18]).

The distance $\bar{d}(\beta^u, \beta^c)$ is calculated based on $E[|h_l(k)|^2]$, which is the long-term power of $h_l(k)$. In the MLSDE receiver, where the CIR is estimated as $\hat{h}_l(k)$, the short-term power of $h_l(k)$ can be approximated by $|\hat{h}_l(k)|^2$ at time k. Based on the short-term power of $\hat{h}_l(k)$ and when the minimum distance error event is encountered, the KL distance for the lth branch between the correct and the incorrect branch metric is defined by

$$d_k^l(\beta^u, \beta^c) = \frac{\overline{|\Delta s_l(k)|^2}|\hat{h}_l(k)|^2}{N_0} - \log\left(1 + \frac{\overline{|\Delta s_l(k)|^2}|\hat{h}_l(k)|^2}{N_0}\right). \tag{29}$$

In the AP method, it is assumed that $h_l(k)$ is zero for calculating the branch metrics (or the branch metrics are fused) when $d_k^l(\beta^u, \beta^c)$ is significantly smaller than $\bar{d}(\beta^u, \beta^c)$, i.e., when

$$d_k^l(\beta^u, \beta^c) \le C_2\bar{d}(\beta^u, \beta^c). \tag{30}$$

The coefficient C_2 is a constant, whose value indicates a tradeoff between the performance and the computational complexity of the receiver. In order to maintain similar performances for the partitioned and unpartitioned receiver, C_2 should be substantially smaller than one. By calculating $\bar{d}(\beta^u, \beta^c)$ and employing some further manipulations; from (30), one can find a threshold for the short-term power of $\hat{h}_l(k)$ to fuse the branch metrics (see Section V).

IV. Implementation

Two points should be considered in the implementation of the ASA algorithm. The first one is the nature of the time-variant computational complexity of the ASA algorithm. Although the average computational complexity of the ASA algorithm is much lower than that of the regular MLSD/MLSDE receiver, in the short-term it may be significantly higher than the average. In order to handle the time-variant complexity of the ASA algorithm using a hardware facility having a fixed computational power, we can use a buffering shift register after the sampler at the receiver. The sampled signal values are buffered in the shift register as a frame and then are processed by a hardware facility with a fixed computational power during a variable-length interval [8].

The other point in the ASA algorithm is the extra computational complexity required for calculating the threshold value and the short-term power of the estimated CIR taps in the context of the AT and AP methods, respectively. Since the maximum Doppler shift is usually much smaller than the transmitted signal's bandwidth, the CIR variation is significantly slower than the sampling rate of the received signal. Therefore, it is not necessary to update the threshold value and short-term power of the CIR taps at each time. Based on the fading rate of the channel, one can select a frame of the received signal, for which the channel variation is small and select a constant threshold value for fusing/diffusing the branch metrics in each frame. In this way the extra computational complexity required by the implementation of the ASA algorithm is negligible especially when the per-survivor processing (PSP) method [20] is used in the MLSDE receiver. Meanwhile for slow fading, the short-term

power of the CIR taps can be estimated in association with a time window of the estimated CIR in order to reduce the effects of estimation errors.

V. Computer Simulations

The bit-error rate (BER) performance and the computational complexity (number of multiplications) of the ASA algorithm have been evaluated by computer simulations for frequency flat and selective fading channels having a normalized Doppler fading rate of $f_dT = 0.01$. The autocorrelation function of the CIR was modeled as (see [8], [21], and [22, ch. 3])

$$\begin{aligned} R_h(l_1, l_2; j) &= E\left[h(l_1, k)h^*(l_2, k-j)\right] \\ &= \sum_{l_1=0}^{L} \exp(-bl_1T_s)J_0(2\pi f_d jT_s)\delta(l_1 - l_2) \\ 0 \le & l_1 \le L, 0 \le l_2 \le L, -\infty \le j \le \infty \end{aligned} \tag{31}$$

where J_0 is the zero-order Bessel function, f_d is the maximum Doppler frequency. The delay rate in (31) was chosen as $b^{-1} = 2T_s$ where T_s is the sample period [8]. Furthermore, $L = 0$ and $L = 2$ in (31) for the flat fading and selective fading channels, respectively, i.e., the channel was simulated with three paths in the frequency selective channel. The impulse response of the transmitter filter was a raised-cosine pulse of

$$g(t) = \operatorname{sinc}\left(\frac{t}{T}\right)\left(\frac{\cos\left(\frac{\nu\pi t}{T}\right)}{1 - \left(\frac{2\nu t}{T}\right)^2}\right) \tag{32}$$

where $\operatorname{sinc}(x) = (\sin(\pi x))/\pi x$, the symbol duration, $T = 1$ and $\nu = 0.35$, while $g(k) = g(t)|_{t=kT_s}$, where $J = T/T_s = 2$. Differentially-encoded quadrature phase shift keying (DQPSK) modulation was chosen for our investigations. The number of states in the trellis diagram is four and 16 for flat fading and selective fading channels, respectively [2]. The Bessel function shaped fading filter modeling an omnidirectional antenna was approximated by an all-pole third-order filter [23]. The number of symbols transmitted in our simulation was $I = 10^5$ in order to ensure that the received signal was faded for adequate durations of time and the data sequence was divided into a sequence of frames of length L_f, where each frame contains 160 data DQPSK symbols ($L_f = 160$). We apply the ASA algorithm in the regular MLSDE receiver, where a Kalman-filter-based channel estimation with known channel parameters is used [2] and compare the BER performance and the computational complexity of the ASA algorithm with the corresponding characteristics of this MLSDE type receiver.

Fig. 7 shows the BER performance of the regular MLSDE and the ASA algorithm, when only the AT method is used for the flat fading channel. The AT value is calculated from (17). For the ASA-AT method, C_1 is chosen to be a fixed value for all signal-to-noise ratios (SNRs) such that the probability of removing the correct state becomes $p_c^r = 0.001$. However, in the ASA-AT* scheme, a different value of C_1 is chosen for each SNR such that $p_c^r = 0.1 \times 10^{-\mathrm{SNR}/12.5}$. Meanwhile, the threshold value is recomputed only after every ten symbols

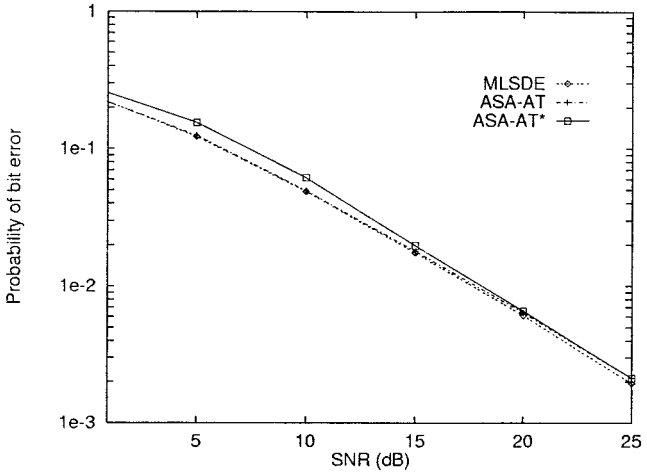


Fig. 7. BER performance of the regular MLSDE and the ASA algorithms for flat fading channel with $f_dT = 0.01$ and DQPSK signaling. The AT method is used in the ASA algorithm.

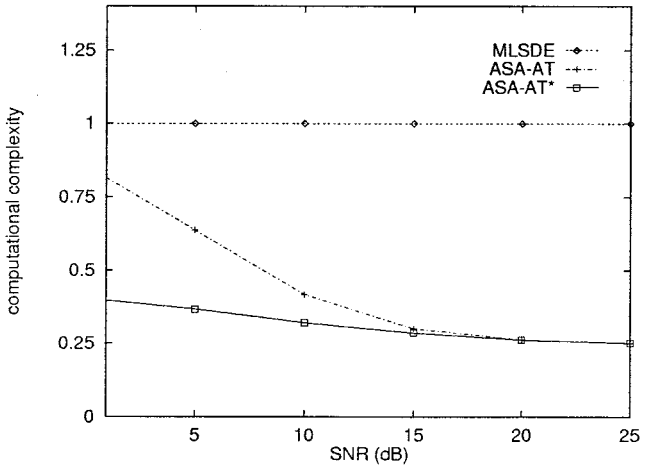


Fig. 8. Computational complexity of the regular MLSDE and the ASA algorithm for flat fading channel with $f_dT = 0.01$ and DQPSK signaling. The computational complexity of algorithms is normalized to that of the regular MLSDE. The AT method is used in the ASA algorithm.

and then it remains constant for the duration of ten symbols, which is sufficiently frequent at the fading rate assumed. As seen in Fig. 7, the performance difference between the regular MLSDE and the ASA-AT algorithms is negligible. However, the performance difference between the regular MLSDE and the ASA-AT* algorithms increases, when the SNR decreases.

As seen in Fig. 8, the computational complexity of the ASA algorithm (both ASA-AT and ASA-AT*) is about 75% lower than that of the regular MLSDE algorithm at $\mathrm{SNR} = 25$ dB. In the context of the AT method, the minimum computational complexity reduction is achieved, when only one trellis state is selected at each time. Since in the flat fading channel model the total number of states is $N_s = 4$, the maximum computational complexity reduction is 75%. As Fig. 8 shows, the proposed AT method achieves a similar level of reduction at $\mathrm{SNR} = 25$ dB. Meanwhile, as Fig. 8 shows, the ASA-AT* method achieves a lower computational complexity in comparison with the ASA-AT method at low SNRs due to selecting a larger p_c^r value in the ASA-AT* method.

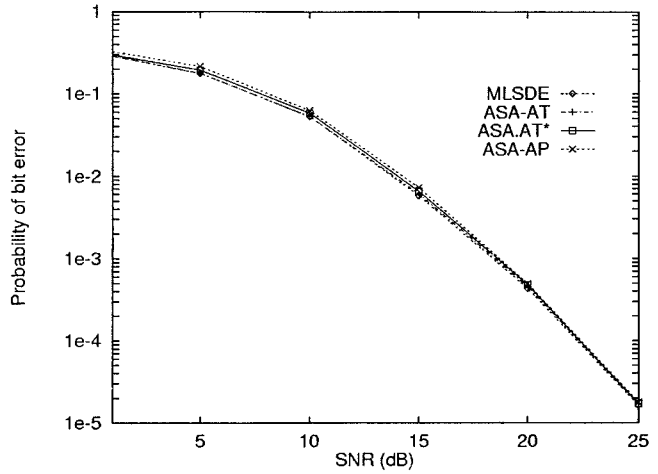


Fig. 9. BER performance of the regular MLSDE and the ASA algorithms for frequency selective fading channel with $f_dT = 0.01$ and DQPSK signaling. The AT method is employed in ASA-AT and ASA-AT* and adaptive state partitioning method is used in ASA-AP.

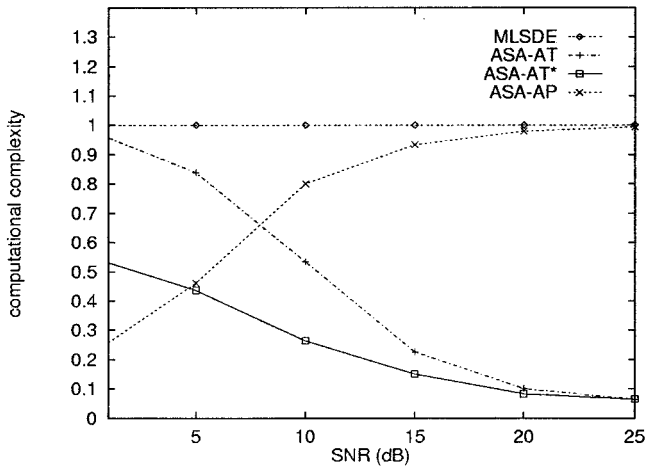


Fig. 10. Computational complexity of the regular MLSDE and the ASA algorithms for frequency selective fading channel with $f_dT = 0.01$ and DQPSK signaling. The computational complexity of algorithms is normalized to that of the regular MLSDE. The AT method is employed in ASA-AT and ASA-AT* and adaptive state partitioning method is used in ASA-AP.

The error rate performances of the regular MLSDE and the ASA algorithms for both AT (ASA-AT and ASA-AT*) and AP (ASA-AP) methods are shown in Fig. 9 for a frequency selective fading channel having three paths, i.e., $L = 2$. According to the AT method, due to the reduced dynamic changes of the CIR tap power associated with the frequency selective channel and as a result of readily obtaining the threshold value, the maximum threshold value, which is constant, is calculated from (19) and C_1 is chosen such that the probability of removing the correct state becomes $p_c^r = 0.001$ and $p_c^r = 0.1 \times 10^{-\mathrm{SNR}/12.5}$ for the ASA-AT and ASA-AT* methods, respectively.

In the AP method, since the sensitivity of the receiver's performance to the difference between the correct and incorrect branch metrics is increased for increasing SNRs, we experimentally chose $C_2 = 0.1 \times 10^{-\mathrm{SNR}/10}$ (C_2 is approximately proportional to the error probability, $C_2 \approx 0.1p(\varepsilon)$) in (30), where $C_2 = 0.1$ and 0.001 for $\mathrm{SNR} = 0$ dB and $\mathrm{SNR} = 20$ dB, respectively. As seen in Fig. 9, the error rate performances of the ASA-AT, ASA-AT* and ASA-AP methods are close to that of the regular MLSDE receiver.

The computational complexities of the regular MLSDE, ASA-AT, ASA-AT*, and ASA-AP methods are shown in Fig. 10. The complexity of the ASA-AT (or ASA-AT*) method is about 93.5% lower than that of the regular MLSDE at $\mathrm{SNR} = 25$ dB and it is increased, when decreasing the SNR. The ASA-AT method achieves the maximum complexity reduction, which is 93.75%, by selecting only one state from 16 states of the trellis diagram where $N_s = 16$. In contrast to the ASA-AT method, the complexity of ASA-AP method is decreased at low SNRs and for high SNRs it is close to the complexity of the regular MLSDE. This phenomenon can be explained in the extreme case, when the noise power is zero or $\mathrm{SNR} = \infty$. In this situation, the correct branch metric is always zero and we should consider all the CIR taps for calculating the branch metrics in order to select the correct path. Therefore, the computational complexity of the ASA-AP method becomes the same as the computational complexity of the regular MLSDE

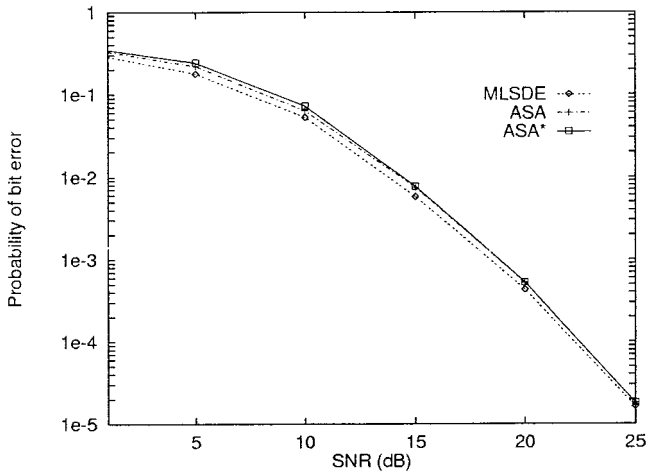


Fig. 11. BER performance of the regular MLSDE and the ASA algorithms for frequency selective fading channel with $f_dT = 0.01$ and DQPSK signaling. Both AT and adaptive state partitioning methods are used in the ASA algorithm.

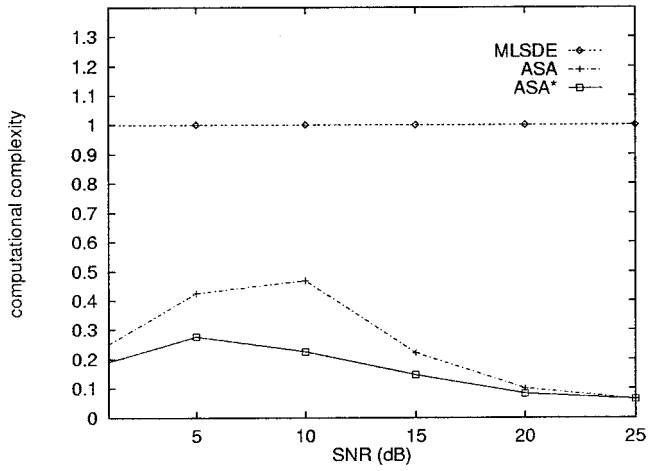


Fig. 12. Computational complexity of the regular MLSDE and the ASA algorithms for frequency selective fading channel with $f_dT = 0.01$ and DQPSK signaling. The computational complexity of algorithms is normalized to that of the regular MLSDE. Both AT and adaptive state partitioning methods are used in the ASA algorithm.

at high SNRs. Meanwhile, although the performance difference between the ASA-AT and ASA-AT* methods is negligible at low SNRs, the difference in their computational complexity becomes significant when the SNR decreases. We believe that due to choosing the maximum threshold value in the selective fading channel, the performance difference between the ASA-AT and ASA-AT* methods becomes negligible in comparison to this difference in flat fading.

Fig. 11 shows the error rate performance of the ASA algorithm when both the AT and AP methods are used. In the ASA algorithm, both the ASA-AT and ASA-AP methods are used at the same time; but in the ASA* algorithm, the ASA-AT* and ASA-AP methods are used together. The performance degradation of the ASA (or ASA*) algorithm is seen to be negligible in comparison to that of the regular MLSDE. Due to the application of both the AT and AP methods, the computational complexity of the ASA (or ASA*) algorithm shown in Fig. 12 is very much lower than that of the regular MLSDE at high and low SNRs. From Fig. 10, one can conclude that for reducing the complexity only at high or low SNRs using the AT method or AP method is sufficient.

VI. Conclusion

We have extended and modified the idea of the ASA algorithm proposed in [8] in order to reduce the computational complexity of the regular MLSDE receiver, when the CIR or its estimation is available at the receiver. Based on the short-term power of the estimated CIR taps (or the actual CIR), only a few trellis states are chosen as the more likely correct states in this algorithm. The proposed ASA algorithm significantly decreases the computational complexity of the regular MLSDE, such that typically only one trellis state (the minimum possible number of states) is chosen at high SNRs without a significant performance loss.

The ASA algorithm is a combination of two techniques, which are the AT and AP methods. In the AT method, which decreases the computational complexity of the regular MLSDE mostly at high SNRs, the threshold value was formulated based on the probability of removing the correct state. The threshold value changes according to the short-term power of the estimated CIR taps (or the actual CIR). The maximum value of the threshold, which can be calculated without needing the CIR estimate, was proposed and can be used for ISI-contaminated channels as well. The AP method, which decreases the computational complexity of the regular MLSDE at low SNRs, was developed based on the KL distance between the probability density functions of the correct and incorrect branch metrics. Therefore, using one of these methods is sufficient when the receiver is designed only for high or low SNRs. However, if the probability of removing the correct state is selected to be larger for low SNRs, the complexity reduction of the AT method will be higher at low SNRs at the expense of a decrease in performance. Meanwhile, although we have proposed the ASA algorithm for MLSD/MLSDE receiver, this algorithm can be applied in symbol by symbol detection as well [12].

References

[1] 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.

[2] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Commun.*, vol. 47, pp. 1181–1193, Aug. 1999.

[3] 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.

[4] 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.

[5] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. COM-36, pp. 13–20, Jan. 1988.

[6] 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.

[7] S. J. Simmons, "Breadth-first trellis decoding with adaptive effort," *IEEE Trans. Commun.*, vol. 38, pp. 3–12, Jan. 1990.

[8] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive state allocation algorithm in mlsd receiver for multipath fading channels: Structure and strategy," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 174–187, Jan. 1999.

[9] N. McGinty and R. Kennedy, "Reduced-state sequence estimator with reverse-time structure," *IEEE Trans. Commun.*, vol. 45, pp. 265–268, March 1997.

[10] J. Belzile and D. Haccoun, "Bidirectional breadth-first algorithms for the decoding of convolutional codes," *IEEE Trans. Commun.*, vol. 41, pp. 370–380, Feb. 1993.

[11] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, pp. 1000–1011, Feb. 1994.

[12] J. P. Seymour and M. P. Fitz, "Near-optimal symbol-by-symbol detection schemes for flat Rayleigh fading," *IEEE Trans. Commun.*, vol. 43, pp. 1525–1533, Feb./Mar./Apr. 1995.

[13] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive T-algorithm in MLSD/MLSDE receivers for fading channels," in *Proc. ICC'99*, vol. I, Vancouver, Canada, June, pp. 539–543.

[14] H. Zamiri-Jafarian, "Adaptive MLSDE Receivers for Wireless Communications," Ph.D., Dept. Elec. Comput. Eng., Univ. Toronto, Toronto, Canada, 1998.

[15] E. A. Lee and D. G. Messerschmitt, *Digital Communications*, 2nd ed. Boston, MA: Kluwer Academic, 1994.

[16] 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.

[17] V. Krishnamurthy and J. B. Moore, "On-line estimation of hidden markov model parameters based on the Kullback-Leibler information measure," *IEEE Trans. Signal Processing*, pp. 2557–2573, Aug. 1993.

[18] V. Krishnamurthy, "On-line estimation of dynamic shock-error models based on the Kullback-Leibler information measure," *IEEE Trans. Automat. Contr.*, pp. 1129–1135, May 1994.

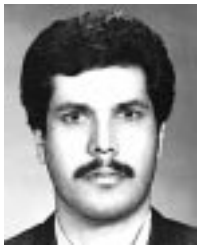
[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[20] R. Raheli and A. Polydoros, "Per-survivor processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, pp. 354–364, Feb./Mar./Apr. 1995.

[21] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, pp. 957–1000, July–Aug. 1968.

[22] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*. New York: Wiley, 1995.

[23] W. H. A. Guston and E. H. Walker, "A multipath fading simulator for radio," *IEEE Trans. Veh. Technol.*, pp. 241–244, Nov 1973.



Hossein Zamiri-Jafarian (M'98) 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, Canada, in 1998, all in Electrical Engineering.

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

Subbarayan Pasupathy (M'73–SM'81–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 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. degree 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 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 at the University of Toronto. His research interests are in the areas of communication theory, digital communications, and statistical signal processing.

Dr. Pasupathy is a registered Professional Engineer in the province of Ontario. During 1982–1989 he was an Editor for Data Communications and Modulation for the IEEE TRANSACTIONS ON COMMUNICATIONS. He has also served as a Technical Associate Editor for the IEEE COMMUNICATIONS MAGAZINE (1979–1982) and as an Associate Editor for the *Canadian Electrical Engineering Journal* (1980–1983). Since 1984, he has been writing a regular column entitled "Light Traffic" for the IEEE COMMUNICATIONS MAGAZINE