Titre: Title:	Fast adaptive sequence estimation over frequency selective fading channels
Auteurs: Authors:	Serge Forest, & David Haccoun
Date:	1994
Type:	Rapport / Report
Référence: Citation:	Forest, S., & Haccoun, D. (1994). Fast adaptive sequence estimation over frequency selective fading channels (Rapport technique n° EPM-RT-94-18). https://publications.polymtl.ca/9878/

Institution:	École Polytechnique de Montréal
Numéro de rapport: Report number:	EPM-RT-94-18
URL officiel: Official URL:	
Mention légale: Legal notice:	

EPM/RT-94/18

Fast Adaptive Sequence Estimation over Frequency Selective Fading Channels

Serge Forest, étudiant M.Sc.A.
Dr. David Haccoun, Professeur Titulaire.

École Polytechnique de Montréal

Mai 1994

Dépôt légal, 2e trimestre 1993
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

Pour se procurer une copie de ce document, s'adresser au:

Éditions de l'École Polytechnique de Montréal
École Polytechnique de Montréal
Case Postale 6079, Succursale Centre-Ville
Montréal (Québec) H3C 3A7
(514) 340-4000

Compter 0,10$ par page (arrondir au dollar le plus près) et ajouter 3,00$ (Canada) pour la couverture, les frais de poste et la manutention. Régler en dollars canadiens par chèque ou mandat poste au nom de l'École Polytechnique de Montréal. Nous n'honorerons que les commandes accompagnées d'un paiement, sauf s'il y a eu entente préalable dans le cas d'établissements d'enseignement, de sociétés ou d'organismes canadiens.

Fast Adaptive Sequence Estimation over Frequency Selective Fading Channels

Serge Forest and David Haccoun

Ecole Polytechnique de Montréal
Department of Electrical and Computer Engineering

May 1994

Fast Adaptive Sequence Estimation over Frequency Selective Fading Channels*

Serge Forest and David Haccoun

Abstract

Sequence estimation is an equalization procedure that may be used for intersymbol interference reduction in frequency selective channels. Like other equalizers, this procedure must be adaptive in order to keep up with channel variations. Unfortunately, conventional adaptive sequence estimators suffer from substantial performance degradations under fast channel variations. Here, a new approach is considered to the adaptive sequence estimation problem in fast fading frequency selective channels. A model for a combined channel and sequence estimation (CCSE) is proposed. This model is used to assess the existing algorithms and to present viable solutions that provide significant improvements in error performances. An analysis of the pairwise distances in the framework of CCSE is also presented. Finally, computer simulation results supporting the theoretical analysis and demonstrating the superiority of the new approach over existing methods in fast fading channels are provided.

* This research has been supported in part by the Natural Sciences and Engineering Research Council of Canada.

I. Introduction

Forney introduced maximum-likelihood sequence estimation (MLSE) in the early 1970s as an efficient way to reduce intersymbol interference (ISI) problems in time-dispersive channels [1]. Since then, this technique has been implemented in a variety of applications including equalization in frequency selective, fast fading channels. MLSE is performed on the premise that the ISI channel can be modeled as a finite state machine. All possible channel state variations are then searched in a maximum-likelihood fashion. When the total number of possible channel states is small, an optimum search procedure (a modified Viterbi algorithm) can be used as a sequence estimator. This type of equalizer can eliminate most of the harmful effects due to intersymbol interference [1].

In order to construct an accurate model of the finite state machine at the receiver, instantaneous channel characteristics must be known at all times. Hence, in a frequency selective fading channel, the sequence estimator must be made adaptive in order to track channel variations. Historically, the first adaptive sequence estimators used decision-directed channel estimators (employing a stochastic gradient algorithm) in conjunction with MLSE [2],[3]. Unfortunately, these estimators have a very slow convergence rate and hence are unable to keep up with fast channel variations. Recently, a number of algorithms designed to improve the convergence and tracking properties of adaptive sequence estimation have been proposed [4]-[7]. As described in some details in section II, all of these algorithms use several channel estimates that are independently adapted to their associated data sequence. For the tracking of channel parameters under fast fading conditions, this approach does not provide a significant improvement of the error performance over the traditional method unless it operates at a fairly high signal-to-noise ratio [5]. Taking into account the significant increase in complexity over the traditional procedure, this approach may not be very interesting from an engineering point of view. This paper presents an analysis of the problems associated with adaptive sequence estimation and proposes practical solutions that provide substantial error performance improvements in frequency selective, fast fading channels. In our approach, the problem is considered as a combined channel and sequence estimation (CCSE) where theoretical models and analysis provide much insight about adaptive sequence estimation, leading to viable solutions.

The paper is organized as follows. Section II presents an overview of existing adaptive sequence estimation techniques. In section III, a model for the combined channel and sequence estimation (CCSE) problem is proposed. Based on that model, a new algorithm is presented together with a theoretical analysis of the pairwise distance properties of CCSE. Results of computer simulations showing that the new algorithm leads to an improvement of the bit error rate by a factor of 5 over the existing methods at $E_b/N_0 = 25\text{dB}$ in certain fast fading channels are provided in section IV and some conclusions are presented in section V.

II. Adaptive sequence estimation

In this section, we review the basic principles of adaptive sequence estimation and present known equalization procedures for frequency selective fading channels.

A. Maximum-likelihood sequence estimation

Let us consider the time dispersive channel model of memory length L shown in figure 1. This model assumes a whitening matched filter is used at the receiver. Each of the $L+1$ tap gain coefficients is independent from the others. Furthermore, in a fading environment, each coefficient

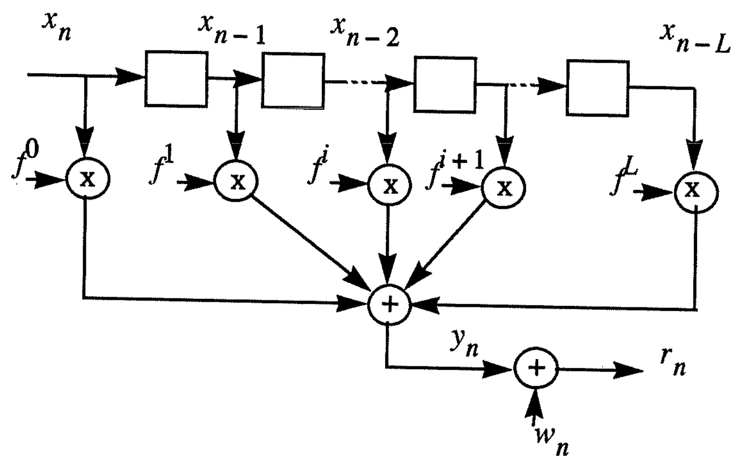

Fig. 1: Model for a time dispersive channel.

has a Rayleigh distributed amplitude and a uniform distributed phase [8]. The complex received signal in the presence of noise at time n is given by

$$r_n = \underline{F}_n^t \underline{X}_n + w_n = y_n + w_n \qquad n = 1, 2, 3, \ldots \tag{1}$$

where y_n, called the filtered data, is the output of the finite state machine, w_n is a zero mean gaussian random variable and where $\underline{X}_n$ and $\underline{F}_n$ are respectively the data symbols vector and the channel tap gain (CTG) vector defined as follows:

$$\underline{X}_n = [x_n, x_{n-1}, \ldots, x_{n-L}]^t \tag{2a}$$

$$\underline{F}_n = [f_n^0, f_n^1, \ldots, f_n^L]^t \tag{2b}$$

where $[\ ]^t$ indicates vector transposition.

The MLSE receiver exploits the channel model of figure 1 by searching among all possible state variations the data sequence that yields the closest filtered data sequence, in Euclidean distance, to the received sequence. The channel state σ_n is defined as the last L entries in the finite state machine:

$$\sigma_n = \{x_n, x_{n-1}, \ldots, x_{n-L+1}\} \tag{3}$$

For a modulation scheme using m signals, there is a total of m^L possible channel states.

The algorithm must select the data sequence $\{\hat{x}\}_i$ of length N symbols that offers the highest cumulative metric:

$$\max_{\{\hat{x}\}_i} \left[\Gamma_{N,i} = -\sum_{n=1}^{N} \left| r_n - \underline{F}_n^t \underline{\hat{X}}_{n,i} \right|^2 = -\sum_{n=1}^{N} \gamma_{n,i} \right] \tag{4}$$

where $\Gamma_{N,i}$ is the cumulative metric for the N symbols associated with the data sequence $\{\hat{x}\}_i$, $\underline{\hat{X}}_{n,i}$ is the associated data symbols vector at time n and where $\gamma_{n,i}$ is the branch metric of path i at time n. The search procedure is usually implemented as a modified Viterbi algorithm if the total number of channel states is small [1].

The major drawback of this procedure is that the calculation of the optimal metric in (4) requires knowledge of the channel tap gain vector $\underline{F}$ at each instant. Since that vector is unknown at the receiver, an estimate $\hat{\underline{F}}$ must be used. In order to make the technique suitable under fast varying channel conditions, one should examine adaptive channel estimators. Some of these adaptive channel estimators used in conjunction with MLSE are presented in the remainder of this section.

B. Conventional adaptive procedure

A decision-directed channel estimator may be used in conjunction with MLSE to make the procedure adaptive. The channel estimates can be continuously adjusted according to a stochastic gradient or recursive least squares (RLS) algorithm [8]. Tentative decisions of the sequence estimator are used to form an error signal which is then employed to make appropriate changes to the current channel estimate. Of course, the efficiency of the procedure is strongly related to the reliability of the tentative decisions. Unfortunately, reliable decisions are not readily available at the output of the MLSE algorithm. A delay δ, called the symbol estimation delay, is necessary between the reception of a given signal and the use of the corresponding decision by the sequence estimator. Using a stochastic gradient algorithm, the channel estimate $\hat{\underline{F}}_n$ at time n is updated as follows:

$$\hat{\underline{F}}_{n+1} = \hat{\underline{F}}_n + \alpha \left(r_{n-\delta} - \hat{\underline{F}}_n^t \hat{\underline{X}}_{n-\delta} \right) \hat{\underline{X}}^*_{n-\delta} \qquad (5)$$

where $r_{n-\delta}$ is the received signal at time $n-\delta$, $\hat{\underline{X}}_{n-\delta}$ is the estimate for the data symbols contained in the finite state machine at time $n-\delta$ (but available only at time n), $(\ )^*$ indicates complex conjugate and where α is the step-size parameter. The paths metrics are computed according to (4), but use the channel estimates:

$$\hat{\Gamma}_{N,i} = -\sum_{n=1}^{N} \left| r_n - \hat{\underline{F}}_n^t \hat{\underline{X}}_{n,i} \right|^2 = -\sum_{n=1}^{N} \hat{\gamma}_{n,i} \qquad (6)$$

where $\hat{\Gamma}_{N,i}$ and $\hat{\gamma}_{n,i}$ are the cumulative and branch metrics of the path i, using the channel estimate $\hat{\underline{F}}_n$. The overall procedure is depicted in figure 2.

There is an inherent compromise in the choice of the symbol estimation delay δ. A short delay permits a faster response to channel variations while a long delay insures a greater reliability in symbol estimation, resulting in a better long term convergence. In a fast fading channel, a short

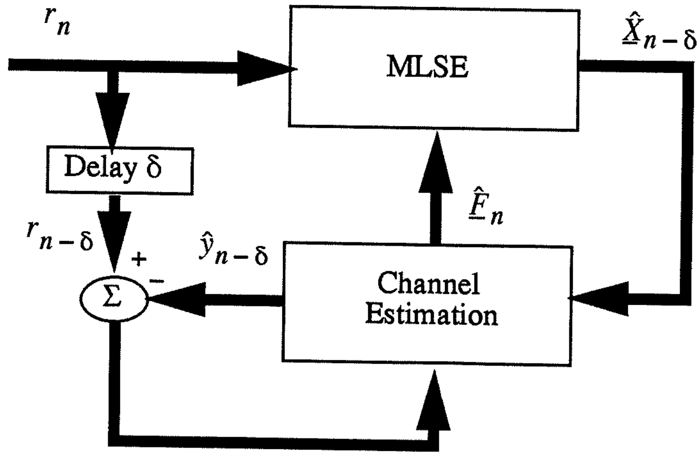

Fig. 2: Conventional adaptive procedure.

delay is essential to keep up with channel variations. A poor reliability of symbol estimates and a substantial decrease in the overall performance are thus unavoidable.

C. Parallel channel estimates

By considering as many independent channel estimates as there are surviving data sequences at each instant, it is possible to shorten the symbol estimation delay while maintaining a certain reliability of the symbol estimates. Each channel estimate is independently adjusted in a decision-directed manner according to their associated data sequence. A number of algorithms using this technique have been recently proposed [4]-[7]. A variety of adaptive procedures were used to continually adjust the channel estimates associated with a particular data sequence [5]. Since recursive least squares (RLS) algorithms do not provide a clear performance improvement over the gradient algorithm for the tracking of fast fading channels and taking into account the relatively high computational complexity of RLS procedures, the gradient algorithm is preferred for the applications considered in this paper.

Using a modified Viterbi algorithm as a sequence estimator in conjunction with the parallel channel estimates technique, each one of the m^L surviving sequences at each instant (m is the number of signals in the constellation and L is the channel memory) must have its own channel estimate based on the tentative decisions contained in its path history. Using a stochastic gradient algorithm, the channel estimate at time n associated with the data sequence $\{\hat{x}\}_i$, $\hat{\underline{F}}_{n,i}$, is updated as follows:

$$\hat{\underline{F}}_{n+1,i} = \hat{\underline{F}}_{n,i} + \alpha (r_n - \hat{\underline{F}}^t_{n,i} \hat{\underline{X}}_{n,i}) \hat{\underline{X}}^*_{n,i} \tag{7}$$

where r_n is the received signal, $\hat{\underline{X}}_{n,i}$ is the column vector of the last $L+1$ symbols contained in the path history i at time n and where α is the step-size parameter. For each extension, the path metric is computed as in (4) but uses the associated channel estimate as follows:

$$\hat{\Gamma}_{N,i} = -\sum_{n=1}^{N} \left| r_n - \hat{\underline{F}}^t_{n,i} \hat{\underline{X}}_{n,i} \right|^2 = -\sum_{n=1}^{N} \hat{\gamma}_{n,i} \tag{8}$$

which is essentially the same as (6), except that in this case each branch metric is evaluated using its own channel estimate.

Once the cumulative metrics have been computed for every possible path extension, the best paths merging at each channel state are selected as survivors. Then, the associated channel estimates are updated according to (7). Thus, the symbol estimation delay is reduced to zero while the uncertainty about this estimation is compensated by considering many possibilities for the transmitted sequence at that time. We refer to this algorithm as the Viterbi sequence estimator with parallel channel estimates (VSE-PCE).

This technique does not however offer a significant performance improvement over conventional methods for the tracking of fast fading channels unless it operates at a fairly high signal-to-noise ratio (SNR>25dB) [5]. Considering the important increase in computational complexity and in memory requirements, this approach seems to have very limited advantages over the conventional adaptive procedure. In the next section, we address the problem of adaptive sequence estimation with a model for a combined channel and sequence estimation. This model is used to assess the above algorithm and to propose a new technique that provides substantial improvement of the overall error performance in frequency selective fading channels.

III. Combined channel and sequence estimation

Every sequence estimator that continuously adjusts its channel estimates, such as the two procedures described in section II, can be considered as an algorithm performing some form of combined channel and sequence estimation (CCSE). But a maximum-likelihood CCSE would require the determination of both the estimated data sequence $\{\hat{x}\}_i$ and the estimated sequence of channel tap gain (CTG) vectors $\{\hat{\underline{F}}\}_j$ which satisfy the following criterion:

$$\max_{\{\hat{x}\}_i,\ \{\hat{\underline{F}}\}_j} P[\{r\} \mid \{x\} = \{\hat{x}\}_i,\ \{\underline{F}\} = \{\hat{\underline{F}}\}_j] \tag{9}$$

where $\{r\}$ is the received signal sequence. An exhaustive search considering all possible data sequences and CTG vector sequences is of course prohibited due to its complexity. In fact, the number of CTG vector sequences alone is infinite if we consider that any channel realization is possible. Below, a simplified search criterion, closer to implementation requirements, is presented.

A. Model of the restricted combined estimation

Let us assume that the available adaptive algorithm can yield only one estimate of the CTG vector sequence, given a specific received signal sequence and a data sequence used as a reference to construct the error signal. This is a reasonable assumption for the RLS algorithms using a fixed forgetting factor and the gradient algorithm using a fixed step-size parameter [8]. Hence, the restriction imposed by the adaptive algorithm can be used to limit the search according to (9). Data sequence estimates and CTG vector sequence estimates are no longer independent. We associate each possible data sequence to the CTG vector sequence that would be delivered by the adaptive algorithm using this data sequence as a reference [4]. For a given channel realization and a received signal sequence, and considering data sequences of length N symbols and a modulation scheme using m signals, the search is thus performed on m^N possible pairs of data/channel estimates. In the restricted CCSE, the algorithm must deliver the pair of data/channel estimates $\{\hat{x}, \hat{\underline{F}}\}_i$ which satisfies the following rule:

$$\max_{\{\hat{x},\hat{F}\}_i} \quad P[\{r\} | \{x,F\} = \{\hat{x},\hat{F}\}_i] \tag{10}$$

The optimal restricted CCSE algorithm delivers the $\{\hat{x},\hat{F}\}$ pair having the best cumulative metric, as calculated in (8), among the m^N possibilities. The data structure for the restricted CCSE may be represented by a tree, as shown in figure 3 for binary transmission ($m = 2$) and a two-symbol memory channel ($L = 2$). In figure 3, $\hat{F}_{n,i}$ is the channel estimate associated with the path i at time n and the channel state is given under each branch. An upper branch corresponds to a bit 0 in the data sequence at that position while a lower branch corresponds to a bit 1.

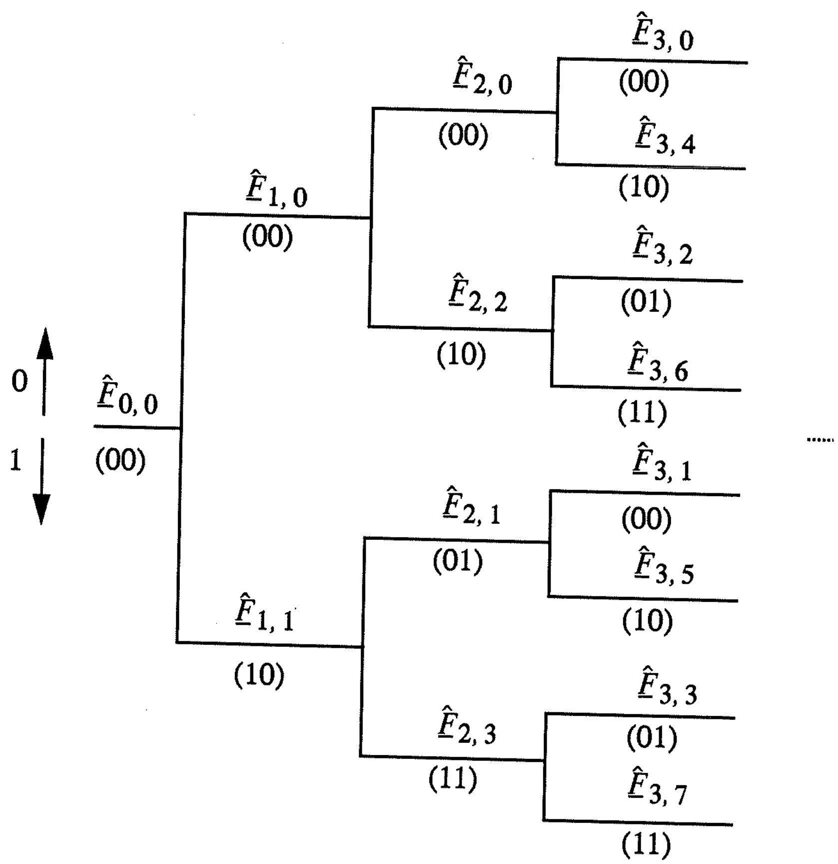

Fig. 3: Restricted CCSE data structure.

Exhaustive exploration of this data structure is obviously very complex even for relatively small sequences. But many sub-optimal search procedures may be considered for the restricted combined estimation. In order to keep a constant computational complexity and delay, breadth-first search algorithms having a fixed number of surviving paths at each instant are considered. One way to explore the data structure depicted in figure 3 is to exploit the merging properties of channel states. As in section II-C, a Viterbi algorithm using parallel channel estimates can be used to perform a sub-optimal search to find the pair $\{\hat{x}, \hat{F}\}_i$ satisfying (10). By doing so, only the m^L best paths merging into *different* channel states are selected at each stage. This selection is made regardless of its impact on the estimation of the channel.

From the above discussion, we see that the major drawback of the VSE-PCE technique presented in section II-C is that the selection of the surviving paths does not depend on the path metrics only. The procedure is also constrained by the selection of paths merging into different channel states. In the case of perfect channel knowledge, this aspect of the selection procedure decreases the number of paths to be explored at no cost to the error performance. However, in the context of restricted CCSE, such a selection rule impairs the algorithm as it forces an additional constraint on the selection of the surviving paths. To avoid this problem, Seshadri [4] proposed to increase the receiver complexity by accepting many survivors per channel state. This approach removes some (but not all) of the constraints associated with the VSE-PCE algorithm at the cost of an important increase in computational load and memory requirements.

We propose to remove all the constraints due to VSE-PCE and to keep an equivalent computational load by using the M-algorithm, also called M-path algorithm, originally introduced in the context of source coding [9]. The M-algorithm has been shown to produce better performance than constrained sub-optimal algorithms having an equivalent complexity [10]. Applied to the restricted CCSE, this algorithm simply selects at each instant the M paths having the highest cumulative metrics, independently of their channel states, and discard the others. By choosing a fixed number of survivors $M = m^L$, the overall algorithm keeps a roughly equivalent computational load as the VSE-PCE [11]. One way to implement this restricted CCSE procedure is to use an M-path sequence estimation in conjunction with parallel channel estimates, which is referred in the following as MSE-PCE. The algorithm can be described, at time n, as follows:

Algorithm

1. Make all data extensions from all surviving pairs at time $n-1$ and compute their metrics according to (8).
2. Select the M pairs having the highest metrics.
3. Update all the associated channel estimates according to (7).

Here, the CCSE model has been used to assess the existing techniques and to propose a new adaptive sequence estimator. In section IV, we present computer simulation results showing that this algorithm outperforms both the traditional method and the VSE-PCE technique. In the following, we utilize the CCSE model to determine the error performance properties of all adaptive sequence estimators using an analysis of the pairwise distances.

B. Pairwise distance properties

The pairwise distances provide much insight about the error performance in the presence of white noise since they allow the determination of the pairwise error probabilities. These probabilities can then be used to upper bound the overall error performance [8]. Here, we focus on the determination of the $\{\hat{x}, \hat{F}\}$ pair distance properties as perceived by a receiver in the context of CCSE. Let us define $D^2[\{\hat{x}, \hat{F}\}_i]$ as the Euclidean distance between a given pair $\{\hat{x}, \hat{F}\}_i$ in the CCSE data structure and the pair corresponding to the correct data sequence. It may be evaluated as the distance between the filtered data sequence $\{\hat{y}\}_i$ corresponding to the pair $\{\hat{x}, \hat{F}\}_i$ and the filtered data sequence $\{\hat{y}\}_{opt}$ corresponding to the correct data sequence and the CTG vector sequence $\{\hat{F}\}_{opt}$ associated to that path. It is thus given by:

$$D^2[\{\hat{x},\hat{F}\}_i] = \sum_{n=1}^{N} |\hat{y}_{n,opt} - \hat{y}_{n,i}|^2$$

$$= \sum_{n=1}^{N} \left|\hat{\underline{F}}^t_{n,opt}\underline{X}_n - \hat{\underline{F}}^t_{n,i}\hat{\underline{X}}_{n,i}\right|^2 \tag{11}$$

where the subscript opt is associated with the $\{\hat{x},\hat{F}\}$ pair corresponding to the correct data sequence. To develop this equation further, it is convenient to define the following vectors called respectively the data error vector $\underline{\varepsilon}$ and the channel error vector $\underline{\xi}$ as:

$$\underline{\varepsilon}_i = \underline{X} - \hat{\underline{X}}_i \tag{12a}$$

and

$$\underline{\xi}_i = \underline{F} - \hat{\underline{F}}_i \tag{12b}$$

For the pair $\{\hat{x},\hat{F}\}_{opt}$ corresponding to the correct data sequence, we have $\underline{\varepsilon}_{opt} = \underline{o}$ and $\underline{\xi}_{opt} = \underline{F} - \hat{\underline{F}}_{opt}$. The channel error vector $\underline{\xi}_{opt}$ is the smallest achievable vector $\underline{\xi}$ with the given adaptive procedure since it corresponds to the error made using the correct reference. Substituting (12a) and (12b) in (11), we obtain

$$D^2[\{\hat{x},\hat{F}\}_i] = \sum_{n=1}^{N} \left|\underline{\xi}^t_{n,i}\underline{X}_n - \underline{\xi}^t_{n,i}\underline{\varepsilon}_{n,i} - \underline{\xi}^t_{n,opt}\underline{X}_n + \underline{F}^t_n\underline{\varepsilon}_{n,i}\right|^2 \tag{13}$$

where the subscript n is the time index and where the subscripts opt and i correspond respectively to the pair of reference $\{\hat{x},\hat{F}\}_{opt}$ and the given pair $\{\hat{x},\hat{F}\}_i$. Using the properties $|a|^2 = a \cdot a^*$ and $2Re\{ab^*\} = ab^* + a^*b$, (13) may be written, after a few algebraic manipulations, as [11]:

$$D^2[\{\hat{x},\hat{F}\}_i] = D^2[\{\hat{x}\}_i] + \wp[\{\hat{x},\hat{F}\}_i] \tag{14}$$

where $D^2[\{\hat{x}\}_i]$ is the data sequence Euclidean distance and where $\wp[\{\hat{x},\hat{F}\}_i]$ is an error term called the tracking perturbation. These terms are given by

$$D^2[\{\hat{x}\}_i] = \sum_{n=1}^{N} \left|\underline{F}^t_n \underline{\varepsilon}_{n,i}\right|^2 \tag{15a}$$

and

$$\wp[\{\hat{x},\hat{\underline{F}}\}_i] = \sum_{n=1}^{N} \Big(\left|\underline{\xi}^t_{n,i}\underline{X}_n - \underline{\xi}^t_{n,i}\underline{\varepsilon}_{n,i} - \underline{\xi}^t_{n,opt}\underline{X}_n\right|^2 + 2Re\{ (\underline{F}^t_n\underline{\varepsilon}_{n,i}) (\underline{\xi}^t_{n,i}\underline{X}_n - \underline{\xi}^t_{n,i}\underline{\varepsilon}_{n,i} - \underline{\xi}^t_{n,opt}\underline{X}_n)^* \} \Big) \tag{15b}$$

Equations (14), (15a) and (15b) give the expression of the distance between the pair $\{\hat{x}, \hat{\underline{F}}\}_i$ and the reference $\{\hat{x}, \hat{\underline{F}}\}_{opt}$ as a function of the CTG vectors, the transmitted data symbols and the error vectors. In the above equations, the effect of the error on the channel estimation has been carefully isolated from the perfect channel knowledge pairwise distance, showing the two factors, (15a) and (15b), that influence the $\{\hat{x}, \hat{\underline{F}}\}$ pairwise distances: the data sequence distances and the tracking perturbation. The former correspond to the pairwise distances obtained with perfect channel state information (CSI) while the latter is the effect of imperfect channel estimates. The two factors combine to yield the total distance.

It is interesting to note the effect of imperfect channel estimates on the pairwise distances. From (15b), we note that the tracking perturbation is not necessarily proportional to the channel error vector $\underline{\xi}$. Hence, a path may have a large error vector $\underline{\xi}$ and still have a small perturbation on its distance from the correct sequence or *vice versa*. Furthermore, for a given channel realization and a transmitted data sequence, the tracking perturbation may be either constructive in the sense that the pair distance $D^2[\{\hat{x}, \hat{\underline{F}}\}_i]$ is larger with that perturbation than with perfect CSI, that is $\wp[\{\hat{x}, \hat{\underline{F}}\}_i] > 0$, or may be destructive in the sense that it reduces the distance as compared with the perfect CSI reception, that is $\wp[\{\hat{x}, \hat{\underline{F}}\}_i] < 0$. In the perfect CSI case, the pairwise data sequence distances determine, for a certain channel realization $\{\underline{F}\}$, the overall error performance. In a CCSE receiver, each one of these pairwise distances is differently modified by the tracking perturbation, sometimes for the better, sometimes for the worse. Only one large destructive perturbation is needed to cause an error event that would not have taken place with an ideal receiver (perfect CSI).

A receiver that has perfect channel knowledge directly improves its error performance with an increase in transmitting power, because of the increase in pairwise distances in (15a). In a CCSE receiver, the effect is somewhat mitigated. A rise in transmitting power is reflected as an increase in data sequence distances because of larger data error vectors but, unfortunately, it is also reflected in higher channel estimation errors on incorrect paths (i.e. increase in $E\left[\underline{\xi}_{n,i}^{t}\,\underline{\xi}_{n,i}^{*}\right]$). Thus, for the same set of CTG vectors and the same transmitted sequence, the variance of the tracking perturbation increases, increasing with it the proportion of error events due to imperfect channel estimates. At high SNR, the tracking perturbation becomes the dominant factor and limits the improvement of the error performance as the transmitting power increases. At a certain point, a greater signal power does not systematically produce an increase in pairwise $\{\hat{x}, \hat{\underline{F}}\}$ distances for a given channel realization. This important consideration should be kept in mind when designing a communication system using adaptive sequence estimation. Note finally that all the sub-optimal algorithms (such as the MSE-PCE and the VSE-PCE) are even more sensible to these phenomena, since they make premature decisions by eliminating certain paths at each instant.

IV. Computer simulations

Several computer simulations were performed in order to assess the algorithms and the analysis presented in the preceding section. The system model (presented below) was not chosen to be especially realistic for a wide-band mobile-radio system, but rather to provide insight about the tracking of fast fading channel parameters and to verify the assertions made in the previous section. The model is described and simulation results are provided in this section.

A. System model

The channel model described in section II was used, with independent Rayleigh fading on each tap. The fading rate was determined by the maximum normalized Doppler frequency $f_d T$, where T is the symbol duration and f_d is the Doppler shift given by

$$f_d = \frac{v}{\lambda} = \frac{v f_c}{c} \qquad (16)$$

where v is the speed of the mobile, λ is the wavelength, c is the speed of light and where f_c is the carrier frequency. We present results for a system having a net transmission rate of 200 Kb/s, using a QPSK modulation and a carrier frequency of 1.2 GHz. Each data sequence contains 500 information bits ($N = 250$ channel symbols). In order to compare the different algorithms in their idealized versions, we have neglected to incorporate all impairments due to mobile-radio systems (other than the ISI problem): imperfect training of the coefficients at the beginning of each sequence, adjacent channel and co-channel interference, phase jitter, etc. In particular, we assumed perfect CSI at the beginning of each sequence and a perfectly coherent receiver.

One channel has played an important role in the computer simulations that were run. This channel uses the "mountainous terrain multipath profile", in which three main paths are present [12]. The second path has a 5dB attenuation compared to the first path and the third path has a 15dB attenuation. Thus, we used a two-symbol memory channel ($L = 2$) with the following mean coefficients: $\overline{|f^0|} = 0.861$, $\overline{|f^1|} = 0.484$ and $\overline{|f^2|} = 0.153$. Other channels, with different impulse responses were also used for the simulations.

The gradient algorithm was used for each adaptive sequence estimator and the step-size parameter was optimized in each situation. The symbol estimation delay δ for the conventional adaptive procedure was also optimized in each channel to produce the best error performance. Finally, all adaptive sequence estimation algorithms were compared to an idealized procedure: a maximum-likelihood sequence estimator with perfect CSI.

B. Simulation results

The first simulations were performed on the mountainous terrain channel model presented above. Figures 4, 5 and 6 illustrate the performances of different algorithms for $f_d T = 5 \times 10^{-4}$, $f_d T = 1 \times 10^{-3}$ and $f_d T = 2.5 \times 10^{-3}$ respectively, which correspond to maximum Doppler shifts of 50 Hz, 100 Hz and 250 Hz in the system described above. Four curves are presented in each figure, which correspond to the following algorithms: the Viterbi sequence estimator with the conventional adaptive procedure, the VSE-PCE algorithm, the MSE-PCE algorithm proposed in section III and the idealized procedure with perfect CSI. For each of the adaptive processes, the step-size parameter having a value $\alpha \approx 0.15$ was found to produce the best error performances. The best symbol estimation delay was found to be $\delta = 4$, $\delta = 2$ and $\delta = 0$ respectively in figures 4, 5 and 6. Finally, a QPSK modulation was used and since the channel has a two-symbol memory, there are $4^2 = 16$ channel states. Hence, we chose $M = 16$ for the M-algorithm so that it would have a similar level of complexity to the Viterbi algorithm.

From figures 4, 5 and 6, one can note that the VSE-PCE algorithm does not provide significant error performance improvements over the conventional adaptive procedure. However, the new MSE-PCE algorithm offers substantial gains over both methods, especially at a SNR of 20dB or more. One can also observe that the error performance of all the adaptive sequence estimators tend to level-off at high SNR. This occurs when the tracking perturbation begins to have

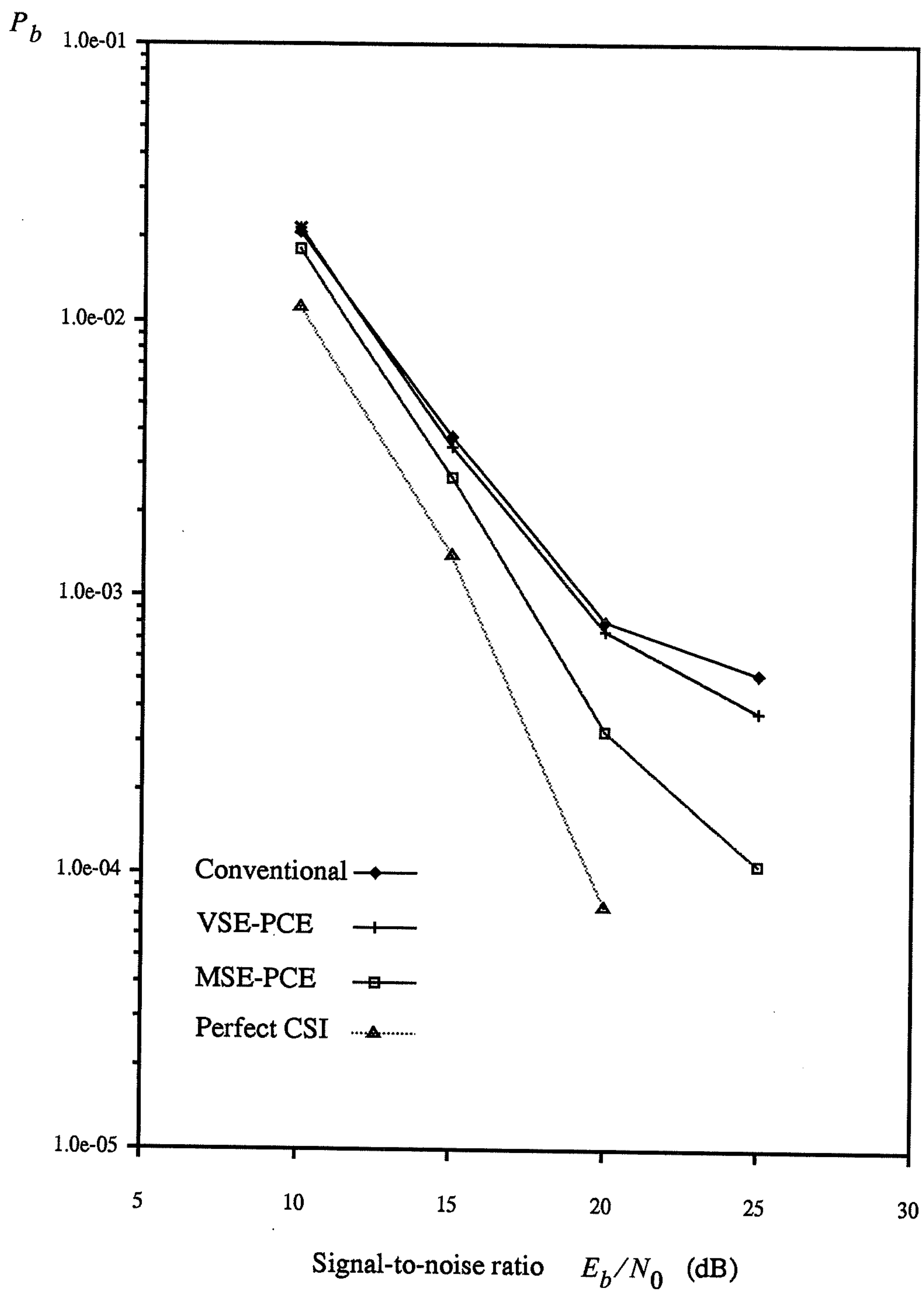

Fig. 4: Bit error probabilities for a 16-state channel with $\overline{|f^0|} = 0.861$, $\overline{|f^1|} = 0.484$ and $\overline{|f^2|} = 0.153$, $f_dT = 5\times10^{-4}$.

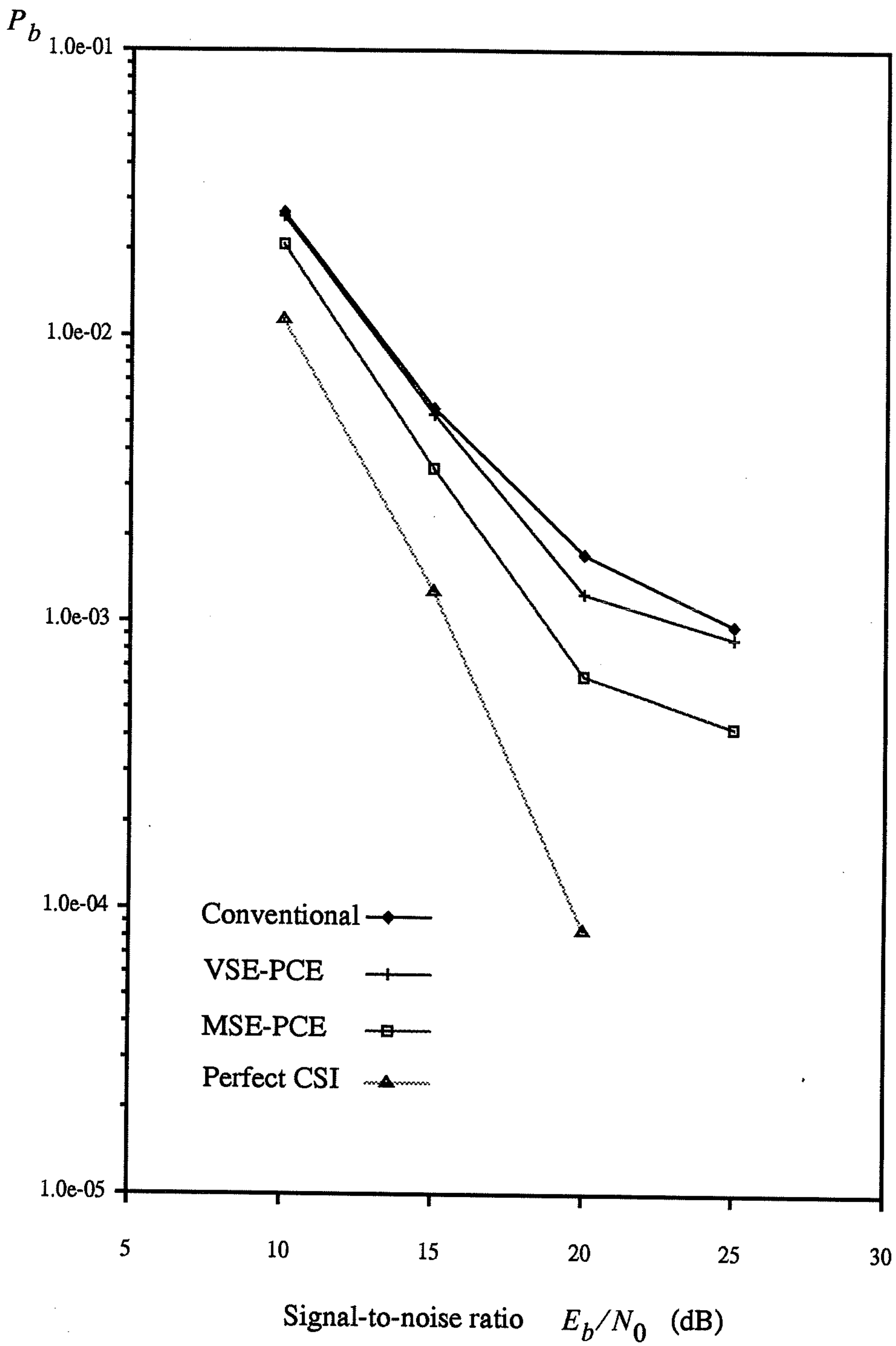

Fig. 5: Bit error probabilities for a 16-state channel with $\overline{|f^0|} = 0.861$, $\overline{|f^1|} = 0.484$ and $\overline{|f^2|} = 0.153$, $f_d T = 1\times10^{-3}$.

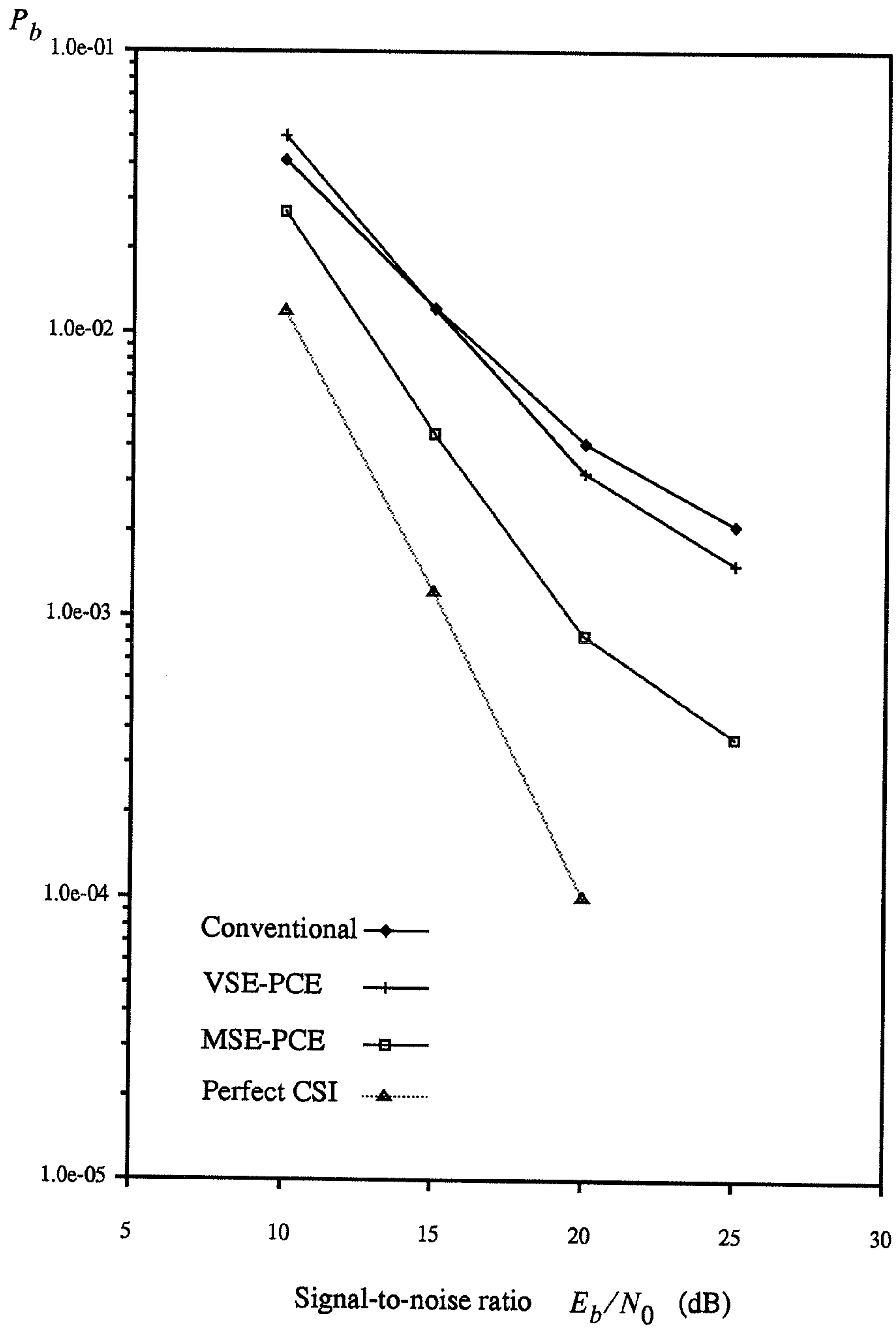

Fig. 6: Bit error probabilities for a 16-state channel with $\overline{|f^0|} = 0.861$, $\overline{|f^1|} = 0.484$ and $\overline{|f^2|} = 0.153$, $f_d T = 2.5\times10^{-3}$.

a noticeable effect. In that respect, the computer simulations tend to support the theoretical assertions made in section III since the analysis predicted that the tracking perturbation would become the dominant factor at high SNR. Indeed, figures 4, 5 and 6 show that the difference between the perfect CSI procedure and the adaptive algorithms increases as the SNR grows. Finally, the improvements provided by the new approach seem to be valid for a wide range of channel variation rates. When the fadings become faster, all the curves corresponding to the three adaptive sequence estimators are shifted away about equally from the performance curve of the idealized procedure.

One of the advantages provided by the M-path algorithm is that its complexity is not directly related to the number of channel states, even though we chose $M = m^L$ in order to compare its computational effort to that of the Viterbi algorithm. In fact, the total computational load can easily be modified by changing the number M of surviving paths at each instant. The effect of M on the error performance for the same channel as above is illustrated in figure 7. We can see that even with half its original computational effort, the MSE-PCE still outperforms both the VSE-PCE algorithm and the conventional adaptive procedure.

A number of simulations were performed to verify that the above results were not specific to the chosen channel. Figure 8 illustrates the performances of the different algorithms in a 16-state and two-symbol memory channel which has the following mean coefficients: $\overline{|f^0|} = \overline{|f^2|} = 0.407$ and $\overline{|f^1|} = 0.815$. This model was used in [6] to characterize the GSM mobile-radio system. In figure 9, the performances of the algorithms are presented for a 64-state channel ($L = 3$) which has the following coefficients: $\overline{|f^0|} = 0.831$, $\overline{|f^1|} = 0.467$, $\overline{|f^2|} = 0.263$ and $\overline{|f^3|} = 0.148$. We used $M = 64$ in that channel for the M-path algorithm to keep its computational load at the same level as the Viterbi algorithm. Finally, figure 10 shows the bit error probabilities in a 4-state channel ($L = 1$) which has the following coefficients: $\overline{|f^0|} = \overline{|f^1|} = 0.707$. This time, the MSE-PCE algorithm used $M = 4$. In all these simulations, we used $f_d T = 1 \times 10^{-3}$ which corresponds to a 100 Hz Doppler shift.

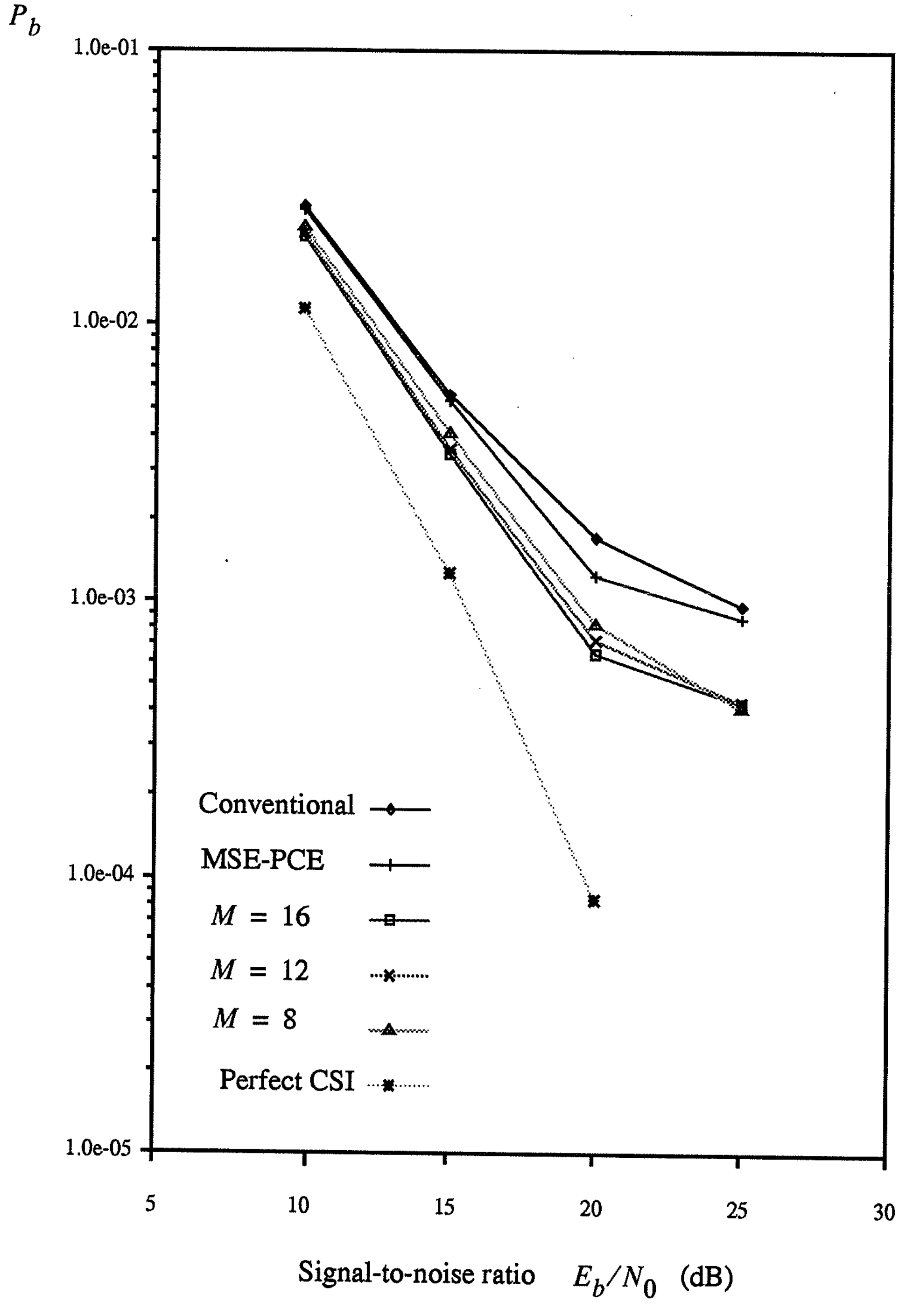

Fig. 7: Bit error probabilities for different values of M in a 16-state channel with $\overline{|f^0|} = 0.861$, $\overline{|f^1|} = 0.484$ and $\overline{|f^2|} = 0.153$, $f_d T = 1\times10^{-3}$.

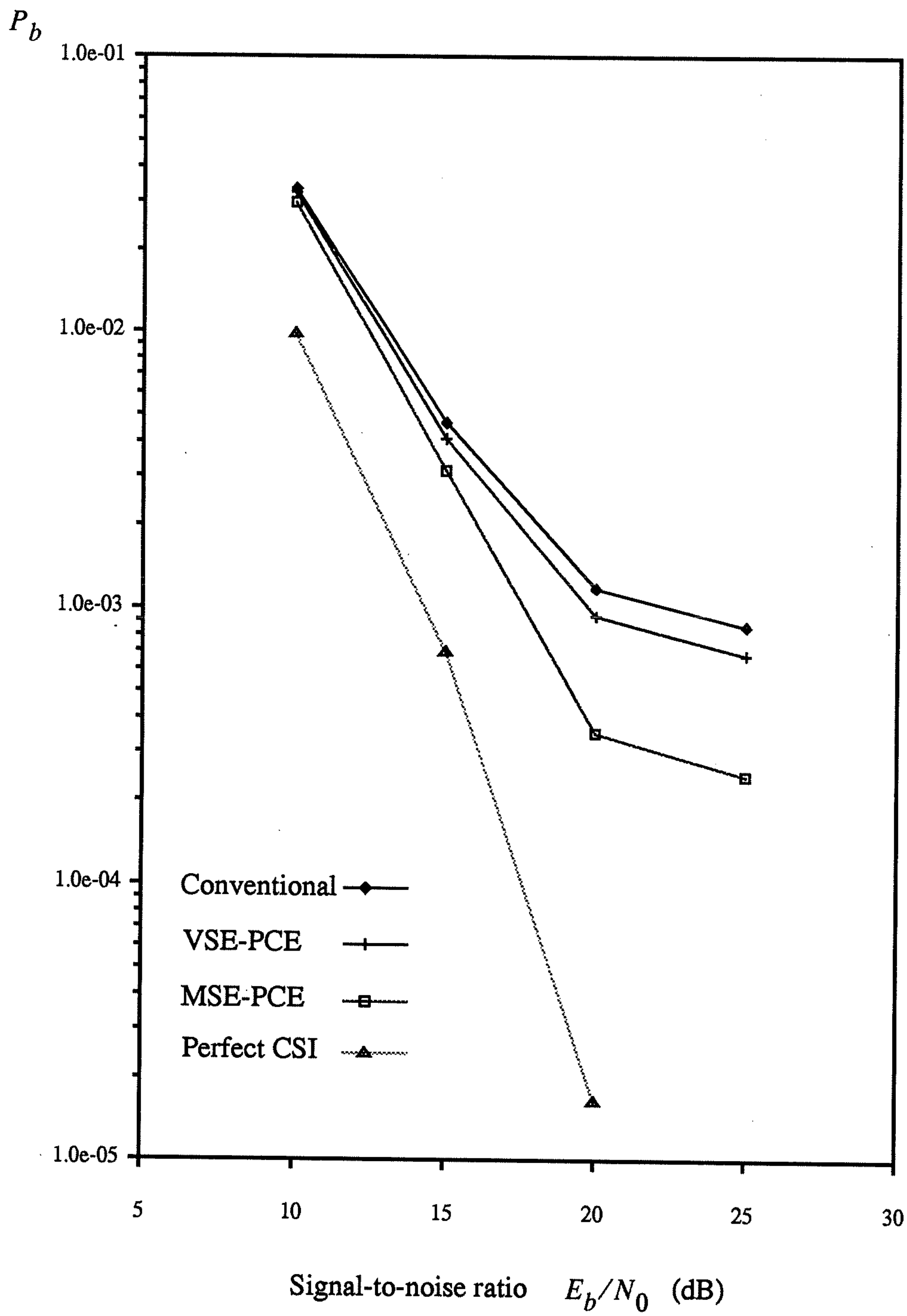

Fig. 8: Bit error probabilities for a 16-state channel with $\overline{|f^0|} = \overline{|f^2|} = 0.407$ and $\overline{|f^1|} = 0.815$, $f_dT = 1\times10^{-3}$.

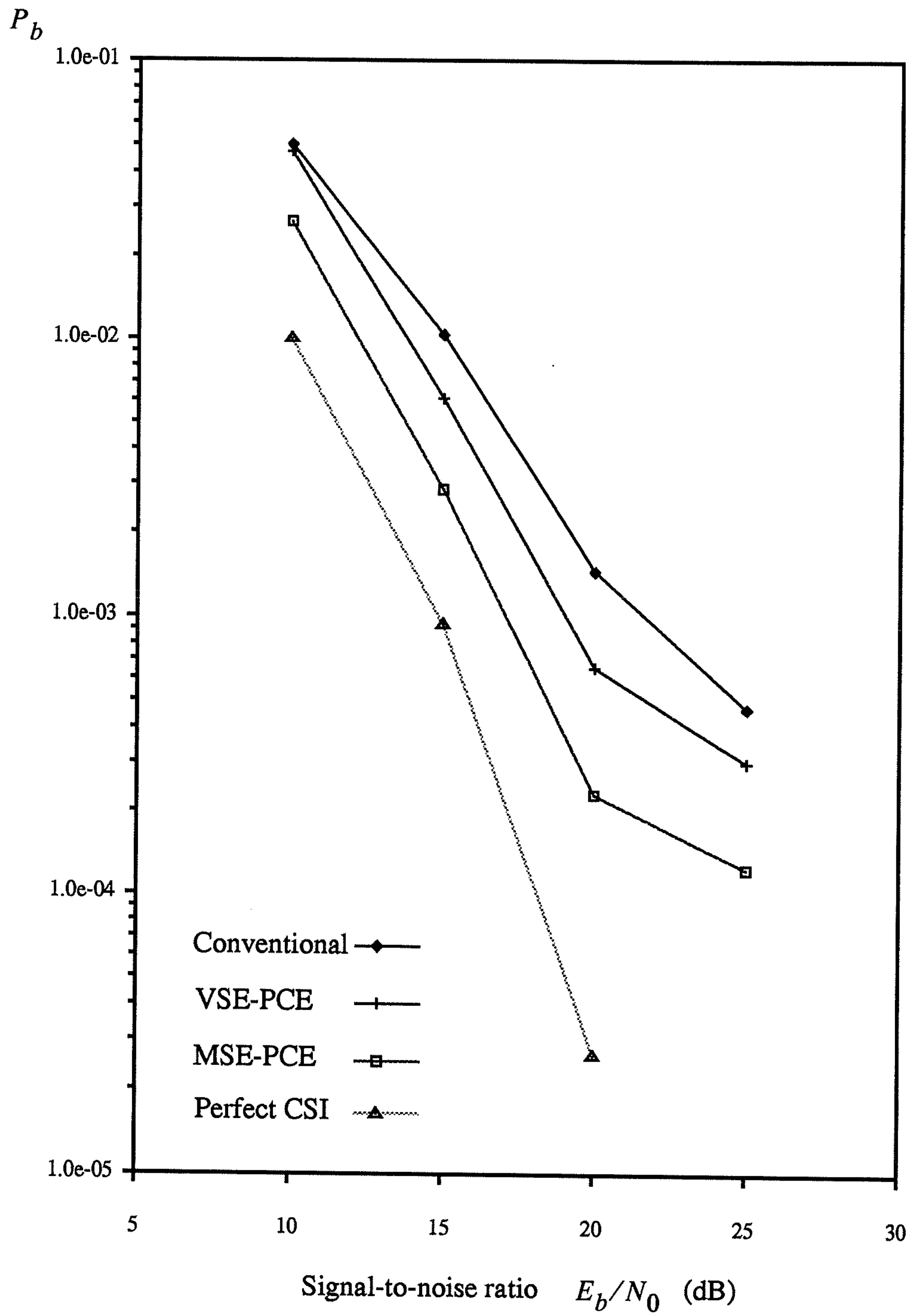

Fig. 9: Bit error probabilities for a 64-state channel with $\overline{|f^0|} = 0.831$, $\overline{|f^1|} = 0.467$, $\overline{|f^2|} = 0.263$ and $\overline{|f^3|} = 0.148$, $f_dT = 1\times10^{-3}$.

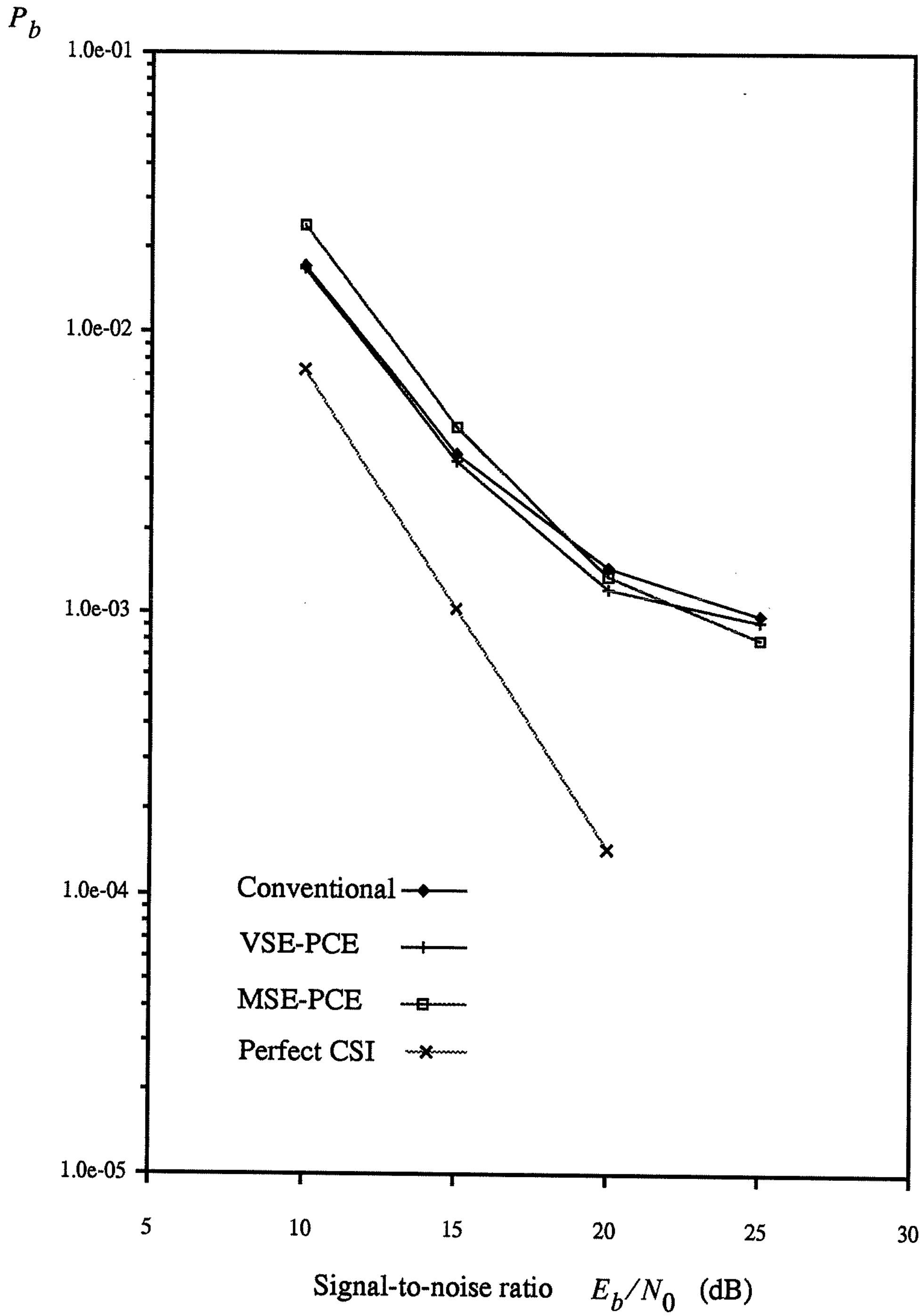

Fig. 10:Bit error probabilities for a 4-state channel with $\overline{|f^0|} = \overline{|f^1|} = 0.707$, $f_d T = 1 \times 10^{-3}$.

Figure 8 shows that the same general trends observed for the previous channel are maintained for the error performances of the different algorithms in that channel. Figure 9 shows that the performances differences between the adaptive algorithms are amplified with an increase of the channel memory. This could be explained by the following facts: the efficiency and flexibility of the M-path algorithm are fully exploited by choosing a large M value since it provides better protection against correct path loss; and the VSE-PCE algorithm has a really clear advantage in computing power for the channel estimation over the conventional adaptive procedure when there is a large number of surviving paths at each instant. These arguments can be reversed to explain why all the procedures have equivalent performances in the one-symbol memory channel, as depicted in figure 10. Simulations performed on other channels having $L = 1$ supported the hypothesis that the new approach does not fully exploit its potential unless it is applied to a channel having $L \geq 2$. An extensive study of all the algorithms presented here can be found in [11].

V. Conclusions

In this paper, we have developed a model for the combined channel and sequence estimation problem. This model has been used to assess the existing algorithms and to develop a new algorithm that fully exploits the principles of the combined estimation. Furthermore, the model has been used to determine error performance properties common to all adaptive sequence estimators with the analysis of the pairwise distances. Thus, two major conclusions may be drawned from this work. First, in adaptive sequence estimation a non-constrained, sub-optimal, equivalent complexity, M-path algorithm used in the combined channel and sequence estimation (CCSE) framework outperforms the traditional Viterbi sequence estimator which is far from optimal in this case. A model for the CCSE problem using a tree data structure has been used to demonstrate that the Viterbi sequence estimator using the parallel channel estimates is limited by the selection of paths merging into different channel states. Computer simulations in fast fading channels have showed that the M-path algorithm, applied to the CCSE of a channel of memory $L \geq 2$ symbols could provide significant error performance improvements over traditional methods, even with reduced computational load. The second major conclusion is that all adaptive sequence estimators have very limited performances at high SNR. In fact, from the two factors influencing the error performance, the data sequence distance properties and the tracking perturbation, the latter becomes dominant as the input power increases. Hence, the tracking perturbation limits the improvement of the error performance as the SNR increases. This consideration must be taken into account in the design of equalizers based on adaptive sequence estimation for fast fading frequency selective channels.

References

[1] G.D. Forney jr., “Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference”, *IEEE Trans. on Inform. Theory,* vol. IT-18, no. 3, May 1972, pp. 363-378.

[2] F.R. Magee jr. and J.G. Proakis, “Adaptive Maximum-Likelihood Sequence Estimation for Digital Signaling in the Presence of Intersymbol Interference”, *IEEE Trans. on Inform. Theory,* vol. IT-19, no. 1, Jan. 1973, pp.120-124.

[3] G. Ungerboeck, “Adaptive Maximum-Likelihood Receiver for Carrier-Modulated Data-Transmission Systems”, *IEEE Trans. Commun.,* vol. COM-22, no. 5, May 1974, pp. 624-636.

[4] N. Seshadri, “Joint Data and Channel Estimation using Fast Blind Trellis Search Techniques”, *IEEE GLOBECOM’90 Conf. Record,* pp.1659-1663.

[5] R. Raheli, A. Polydoros and C.K. Tzou, “The Principle of Per-Survivor Processing: A General Approach to Approximate an Adaptive MLSE”, *IEEE GLOBECOM’91 Conf. Record,* pp. 1170-1175.

[6] K. Fukawa and H. Suzuki, “Adaptive Equalization with RLS-MLSE for Frequency-Selective Fast-Fading Mobile Radio Channels”, *IEEE GLOBECOM’91 Conf. Record,* pp. 548-552.

[7] R.A. Iltis, “A Bayesian Maximum-Likelihood Sequence Estimation Algorithm for a priori Unknown Channels and Symbol Timing”, *IEEE Journ. Select. Areas Commun.,* vol. 10, no. 3, April 1992, pp. 579-588.

[8] J.G. Proakis, *Digital Communications,* 2nd edition, New-York: McGraw-Hill, 1993.

[9] F. Jelinek and J.B. Anderson, “Instrumental Tree Encoding of Information Sources”, *IEEE Trans. on Inform. Theory,* vol. IT-17, Jan. 1971, pp.118-119.

[10] T. Aulin, *An Efficiency Evaluation of Different Simplified Sequence Detection Algorithms*, Technical Report no.106, Department of Telecommunication Theory, Chalmers University of Technology, Göteborg, Sweden, March 1991.

[11] S. Forest, *Estimation Adaptative des Séquences en Présence d'Interférence entre Symboles*, Masters' thesis, Department of Electrical and Computer Engineering, Ecole Polytechnique de Montréal, Montreal, Quebec, May 1994.

[12] R.W. Lorenz, J. deWeck and P. Merki, "Power Delay Profiles Measured in Mountainous Terrain", *Proc. VTC 88,* pp. 105-112.