Sequence Detection in Nonlinear Channels: A Convenient Alternative to Analog Predistortion

Armando Vannucci, *Member, IEEE*, and Riccardo Raheli, *Member, IEEE*

***Abstract*—A new maximum-likelihood sequence detection receiver for spectrally efficient linear modulations on bandlimited bandpass nonlinear channels is proposed. The receiver is based on oversampling the received signal corrupted by noise and nonlinear distortion. Contrary to other solutions in the literature, in the proposed technique there is no need for a bank of matched filters, and the receiver front end reduces to a single lowpass filter. For a given peak power level, a performance gain can be achieved over more traditional approaches to transmission on nonlinear channels, such as those based on predistortion, if a moderate spectral expansion is allowed. To analyze the receiver performance, the concept of *distance spectrum* is employed, since the minimum distance alone cannot account for a reliable performance evaluation. Both analysis and simulation are carried out for realistic narrowband nonlinear channels, possibly employing reduced-state sequence detection. Appreciable gain margins are confirmed to be possible in these realistic cases.**

***Index Terms*—Distance spectrum, maximum-likelihood sequence detection, nonlinear channels.**

I. Introduction

THE analysis and design of digital communication links including nonlinear power amplifiers is a twofold, challenging task. On the practical side, many radio frequency (RF) systems, such as satellite links or terrestrial radio links, force the designer to search for power amplifiers matching the linearity requirements on which the system design is based. This task can involve a dramatic tradeoff that often leads to giving up a fair amount of the available power, for the sake of a linear amplifier transfer characteristic. Moreover, nonlinear phenomena arise also in physical transmission means other than radio channels, like the magnetic and the optical channels, employed for data storage purposes. On the theoretical side, most of the efforts in the literature have been devoted to techniques for eliminating or reducing the nonlinear distortions caused by high power amplifiers, through the use of analog or digital predistortion techniques, including [1]–[5], just to cite a few. Alternative approaches for compensating the nonlinear effects while acting on the receiver were also studied, mainly in the field of nonlinear equalization, among which was the pioneering work presented in [6].

Paper approved by W. E. Ryan, the Editor for Modulation, Coding, and Equalization of the IEEE Communications Society. Manuscript received August 1, 2000; revised July 25, 2001 and February 12, 2002. This work was supported by Siemens ICN, Milan, Italy. This paper was presented in part at the IEEE International Conference on Communications (ICC'98), Atlanta, GA, May 1998, and at the IEEE International Conference on Universal Personal Communications (ICUPC'98), Florence, Italy, October 1998.

The authors are with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy (e-mail: vannucci@tlc.unipr.it).

Publisher Item Identifier 10.1109/TCOMM.2002.802567.

The early work in [7], dealing with performance evaluation, suggests that satellite nonlinearities cause little or no inherent performance degradation which cannot be recovered by an appropriate receiver design. More recent studies, devoted to performance evaluation, state that in some cases the nonlinear performance of an amplifier can be desirable in fading radio channels, the nonlinearity being able to reduce the time dispersion of a linear channel [8]. Based on the idea of exploiting the full available power through an appropriate optimal receiver, there have been some works, among which are [9]–[11], which devise receiver structures for performing maximum-likelihood sequence detection (MLSD). All these solutions share the feature of a remarkable receiver complexity, especially in the analog front end section. Based on a Volterra-series expansion of the nonlinear channel, we devise a novel optimal receiver which makes use of oversampling and the Viterbi algorithm to perform MLSD. Though the joint use of Volterra series and the Viterbi algorithm is not new, see, e.g., [10] and [12], we emphasize that, contrary to these references, the proposed solution uses a much simpler receiver front end with a single filter, instead of a bank of filters. Moreover, the considered solution overcomes some of the practical limitations of previously proposed receivers, and in particular, is characterized by a state complexity which does not depend on the order of the nonlinearity.

The original contribution of this paper is twofold: 1) to propose a new optimal receiver structure operating on nonlinear time-dispersive channels fed by linearly modulated data sequences, while keeping the receiver complexity affordable, especially in the more critical analog front end; 2) to demonstrate, both by analysis and simulation, that the superior potential of nonlinear high power devices can be exploited against noise so that the attainable performance can justify the employment of such a receiver, instead of a less cost- and performance-effective analog predistorter. This paper expands upon preliminary work reported in [13] and [14].

The paper is organized as follows. In Section II, we describe the system model. Section III introduces the new MLSD receiver based on oversampling. Section IV presents an analysis of error performance. Numerical results are given in Section V, and conclusions are drawn in Section VI.

II. System Model

We assume a bandpass nonlinear digital transmission system whose baseband equivalent is shown in Fig. 1. A sequence of independent and identically distributed (i.i.d.) complex symbols $\{a_k\}$ linearly modulate a shaping pulse $p(t)$. The linearly modulated signal $x(t)$ is fed to a nonlinear power amplifier, denoted as NL in the figure, which outputs the nonlinearly distorted signal

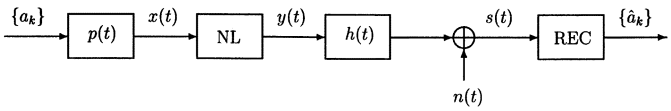

Fig. 1. Baseband equivalent of a nonlinear system.

$y(t)$. This signal exhibits a bandwidth expansion effect caused by the nonlinearity. The power spectrum of $y(t)$ thus depends on the nonlinear characteristic, and may, in general, undergo a bandwidth limitation prior to the transmission on the physical channel. The filter $h(t)$ can be viewed as the cascade of an RF filter that limits the bandwidth of the transmitted signal and a physical channel filter. For brevity, we refer to $h(t)$ as a channel filter. The receiver observes a signal $s(t)$ which is corrupted by additive white Gaussian noise (AWGN) $n(t)$, which is the baseband equivalent of bandpass noise with power spectral density $N_0/2$. The receiver, labeled REC in Fig. 1, will be designed in order to provide a maximum-likelihood estimate $\{\hat{a}_k\}$ of the transmitted sequence.

This system model is of a rather general type, and its main feature is the presence of a single nonlinear memoryless element between a cascade of linear filters. Such systems have been extensively studied and one of the most common approaches to their modeling is through the use of the Volterra kernels [15]. This approach is based on a polynomial expansion of the nonlinear characteristic acting upon the linearly modulated bandpass signal of complex envelope $x(t)$. This polynomial expansion, with coefficients γ_i, is also convenient because it maintains a polynomial form when the nonlinear element is characterized in terms of the baseband equivalents of its input and output signals, as shown in [16]. Referring to Fig. 1, the relationship between $x(t)$ and $y(t)$ can be expressed using a polynomial model of order $2\nu+1$ for the nonlinear element as follows:

$$y(t) = \left[\sum_{m=0}^{\nu} \gamma_{2m+1} \frac{1}{2^{2m}} \binom{2m+1}{m} |x(t)|^{2m}\right] x(t). \quad (1)$$

It is well known that, thanks to the bandpass property, only the odd-order terms of polynomial memoryless bandpass nonlinearities appear in the above expression [12], [16]. In (1), it must be assumed that the coefficients γ_{2m+1} are complex in order to justify both amplitude (AM/AM) and phase (AM/PM) distortions; complex coefficients can be derived, for instance, from the *parallel model* described in [17], which takes into account both nonlinear distortion effects through the use of two real nonlinear elements in two parallel quadrature branches, the second including a phase shift of 90°.

If a linear modulation format, such as M-ary quadrature amplitude modulation (M-QAM), is assumed, the input signal to the nonlinearity takes the general form

$$x(t) = \sum_k a_k p(t - kT) \quad (2)$$

in which T is the symbol interval. We can now rewrite (1) making the dependence on the transmitted sequence explicit; for the sake of simplicity, we assume a third-order, or *cubic*, nonlinearity such that $\nu = 1$

$$\begin{aligned} y(t) =& \gamma_1 \sum_k a_k p(t - kT) \\ &+ \frac{3}{4}\gamma_3 \sum_i \sum_j \sum_l a_i a_j a_l^* p(t - iT) \\ &\cdot p(t - jT) p^*(t - lT) \end{aligned} \quad (3)$$

where a superscript * denotes complex conjugation. This expression can be easily extended to the general case $\nu > 1$ and the derivation that follows can be adapted by adding similar extra terms of higher order. Since for our purposes $\nu = 1$ is sufficient to describe all the relevant concepts, a cubic nonlinearity is assumed in the following.

The signal $y(t)$ is now filtered by the linear system $h(t)$ which, in general, may limit its bandwidth. The received signal is thus

$$\begin{aligned} s(t) =& \gamma_1 \sum_k a_k f(t - kT) + \frac{3}{4}\gamma_3 \sum_i \sum_j \sum_l a_i a_j a_l^* \\ &\cdot \rho(t - iT, t - jT, t - lT) + n(t) \end{aligned} \quad (4)$$

where the following two functions have been used:

$$f(t) \triangleq p(t) \otimes h(t) = \int_{-\infty}^{+\infty} h(\tau) p(t - \tau) d\tau \quad (5)$$

$$\rho(t_1, t_2, t_3) \triangleq \int_{-\infty}^{+\infty} h(\tau) p(t_1 - \tau) p(t_2 - \tau) p^*(t_3 - \tau) d\tau \quad (6)$$

in which $\otimes$ denotes convolution. The first, $f(t)$, is the impulse response of the cascade of shaping and channel filters, or, in general, the cascade of all the linear elements present in the transmission channel. The second function $\rho(t_1, t_2, t_3)$ may be regarded as a third-order time-domain response of the channel: it is indeed closely related to the third-order Volterra kernel of the nonlinear transmission channel. If one wishes to describe the channel through its Volterra kernels as in [15], the first- and third-order kernels are found by simply multiplying $f(t)$ and $\rho(t_1, t_2, t_3)$ by γ_1 and $3\gamma_3/4$, respectively.

III. MLSD Receiver Structure

The bandwidth expansion caused by the nonlinear element, along with the bandlimiting effect of the channel filter $h(t)$, yields a noiseless signal $s'(t) \triangleq s(t) - n(t)$ whose one-sided bandwidth can still exceed the signaling frequency $1/T$. The bandwidth of $y(t)$ is, in general, $2\nu+1$ times that of the shaping pulse $p(t)$.

We assume that the noiseless signal component $s'(t)$ has a one-sided bandwidth B that obeys the following inequality:

$$B < \frac{\sigma}{2T} \quad (7)$$

with σ a properly chosen integer. The signal $s(t)$ is passed through a filter with frequency response $R(f)$ that is nonzero

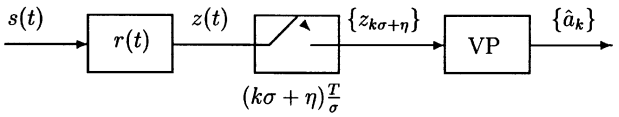

Fig. 2. Optimal receiver structure.

over the bandwidth B and strictly bandlimited to B_R such that the following condition is met:

$$B_R < \frac{\sigma}{T} - B. \tag{8}$$

Based on the results of [18] and [19], and in particular on the *reversibility principle*, the set of samples obtained by sampling the output of this filter with rate σ/T is a sufficient statistic for the detection of the transmitted sequence $\{a_k\}$. We then refer to the receiver structure shown in Fig. 2, in which the only task of the front-end filter with impulse response $r(t)$ is to preserve the useful spectral components of $s(t)$ and limit the bandwidth of the noise component. The samples $z_{k\sigma+\eta}$, at the output of a sampler which operates with period T/σ, are indexed by subscripts k, that scans the symbol period, and $\eta = 0, 1, \ldots, \sigma-1$, that scans the σ samples extracted in the kth symbol period. Provided the "oversampling factor" σ satisfies (8), the "observed" sequence $\{z_{k\sigma+\eta}\}$ is a sufficient statistic and can be used by a suitably designed sequence detector. This sequence detector can be based on the Viterbi algorithm for searching a trellis diagram with proper branch metrics. In Fig. 2, the corresponding Viterbi processor (VP) is shown.

An expression for the observed samples is

$$\begin{aligned} z_{k\sigma+\eta}(a_{k-L}\cdots a_k) =& \gamma_1 \sum_n f^{(\sigma)}_{n\sigma+\eta} a_{k-n} \\ &+ \frac{3}{4}\gamma_3 \sum_i \sum_j \sum_l \rho^{(\sigma)}_{i\sigma+\eta,j\sigma+\eta,l\sigma+\eta} \\ &\cdot a_{k-i}a_{k-j}a^*_{k-l} + w_{k\sigma+\eta} \end{aligned} \tag{9}$$

where the following channel dispersion parameters are defined:

$$f^{(\sigma)}_{n\sigma+\eta} \triangleq f\left(nT + \eta\frac{T}{\sigma}\right) \tag{10}$$

$$\rho^{(\sigma)}_{i\sigma+\eta,j\sigma+\eta,l\sigma+\eta} \triangleq \rho\left(iT+\eta\frac{T}{\sigma}, jT+\eta\frac{T}{\sigma}, lT+\eta\frac{T}{\sigma}\right) \tag{11}$$

and $w_{k\sigma+\eta}$ are filtered noise samples. The parameters in (10) are obtained by oversampling,[1] with rate σ/T, the impulse response of all the cascaded linear elements of the channel: $p(t)$, $h(t)$, and $r(t)$. Such an impulse response is thus the result of the integral in (5), provided that $h(t)$ is replaced by a function $h'(t)$ which results from the cascade of channel and front-end filter: $h'(t) \triangleq h(t) \otimes r(t)$. A similar statement can be made for the third-order parameters in (11), which relate to the function in (6). The reason for using the same symbols, f and ρ, in (5) and (6), as well as in (10) and (11), is that a very simple choice for the front-end filter $r(t)$ can be always made. In fact, by assuming $R(f)$ to be constant in the signal bandwidth B, the front-end filter does not alter the noiseless signal $s'(t)$ and $h'(t) = h(t)$ results. By further assuming $|R(f)|^2$ to have vestigial symmetry around the frequency $\sigma/2T$, uncorrelated noise samples $w_{k\sigma+\eta}$ result. Even though these conditions are not strictly required for an optimal processing, they simplify the sequence detection algorithm. For this purpose, we can resort to a *root-Nyquist* filter with a flat frequency response over the bandwidth B, for which the square modulus $|R(f)|^2$ obeys the first Nyquist criterion for the absence of intersymbol interference (ISI) with signaling frequency σ/T. Among the various possibilities, the extensively used *root-raised-cosine filter* will be assumed for the theoretical study hereafter.

The computation of the channel dispersion parameters in (10) and (11) can be easily carried out by means of the discrete Fourier transform (DFT), noting that the frequency domain equivalents of (5) and (6) are

$$F(f) \triangleq \mathcal{F}[f(t)] = H(f)P(f) \tag{12}$$

$$\begin{aligned} \Gamma(f_1, f_2, f_3) \triangleq \mathcal{F}_3[\rho(t_1, t_2, t_3)] &= H(f_1 + f_2 + f_3)P(f_1) \\ &\cdot P(f_2)P^*(-f_3) \end{aligned} \tag{13}$$

where $\mathcal{F}[\cdot]$ and $\mathcal{F}_3[\cdot]$ denote monodimensional and three-dimensional Fourier transforms, respectively. Replacing $H(f)$ with $H'(f) \triangleq H(f)R(f)$, the last two equations can account for both the channel and receiver filters, whenever necessary. Given the transfer functions of all the filters, an inverse DFT of the above expressions with proper time spacing gives the desired channel dispersion parameters.

Assuming $f(t)$ has strictly finite duration $(L+1)T$, it can be easily seen from (5) and (6) that the result of the integral defining $\rho(t_1, t_2, t_3)$ also vanishes when any of the time variables t_i lies outside an interval of length $(L+1)T$. Thus, all the summations in (9) extend over $L+1$ symbol intervals. Assuming causality, the summation indexes in (9) extend from 0 to L, L being the dispersion length, or memory, of the channel. Clearly, $f(t)$ cannot have finite duration since it is bandwidth limited. If it is approximately so, i.e., if $f(t)$ has only a finite *effective* duration, then it is intuitive that the effective duration of $\rho(t_1, t_2, t_3)$ will also be practically limited to the same interval for any of the time variables t_i. This result was confirmed by numerical computation of channel dispersion parameters for some typical radio channels. We can then conclude that the summations for the nonlinear term in (9) extend over the same interval as the summation which accounts for the linear term.

Maximum-likelihood detection of the transmitted sequence, given the observation of the sequence of samples (9), requires the maximization of the conditional probability density function $p(\mathbf{z}|\mathbf{a})$. The vector of observed samples $\mathbf{z}$, given any particular transmitted sequence $\mathbf{a}$, has a multivariate Gaussian distribution, each sample depending on $L+1$ transmitted symbols and having a conditional expected value equal to the *noiseless part* of that same sample

$$\begin{aligned} E\{z_{k\sigma+\eta}|\mathbf{a}\} &= z'_{k\sigma+\eta}(\mathbf{a}) \\ &\triangleq \gamma_1 \sum_{n=0}^{L} f^{(\sigma)}_{n\sigma+\eta} a_{k-n} \\ &\quad + \frac{3}{4}\gamma_3 \sum_{i=0}^{L}\sum_{j=0}^{L}\sum_{l=0}^{L} \rho^{(\sigma)}_{i\sigma+\eta,j\sigma+\eta,l\sigma+\eta} \\ &\quad \cdot a_{k-i}a_{k-j}a^*_{k-l} \end{aligned} \tag{14}$$

[1]The superscript (σ) in (10) and (11) emphasizes this oversampling.

and a variance given by the variance of the noise samples

$$E\left\{\left(z_{k\sigma+\eta}-z'_{k\sigma+\eta}\right)^2|\mathbf{a}\right\}=\sigma_w^2 \tag{15}$$

with σ_w^2 depending both on noise power spectral density and the receiver filter transfer function. If a receiver filter is chosen according to the aforementioned criteria, these noise samples are uncorrelated, then independent since they are Gaussian. Assuming the transmission of a finite sequence of $I+1$ symbols, the MLSD decision strategy is

$$\hat{\mathbf{a}}=\arg\min_{\mathbf{a}}\left\{\sum_{k=0}^{I+L}\sum_{\eta=0}^{\sigma-1}|z_{k\sigma+\eta}-z'_{k\sigma+\eta}(\mathbf{a})|^2\right\}. \tag{16}$$

Following the celebrated approach in [20], it is now easy to define a recursive sequence metric

$$\begin{aligned}\Lambda_{n+1}(\mathbf{a})&\triangleq\sum_{k=0}^{n}\sum_{\eta=0}^{\sigma-1}|z_{k\sigma+\eta}-z'_{k\sigma+\eta}(\mathbf{a})|^2\\&=\Lambda_n(\mathbf{a})+\lambda(\mu_n\rightarrow\mu_{n+1})\end{aligned} \tag{17}$$

with incremental *branch* metrics

$$\lambda(\mu_n\rightarrow\mu_{n+1})\triangleq\sum_{\eta=0}^{\sigma-1}|z_{n\sigma+\eta}-z'_{n\sigma+\eta}(a_{n-L}\cdots a_n)|^2 \tag{18}$$

defined for each possible *state transition* $(\mu_n\rightarrow\mu_{n+1})$ at the nth symbol interval, where the state at time n is defined in terms of the L previously transmitted symbols $\mu_n\triangleq(a_{n-L},\ldots,a_{n-1})$. The search for the sequence satisfying (16) can now be performed by means of the Viterbi algorithm that traces the evolution of the path metrics in (17) on a trellis diagram with M^L states. The only difference with respect to the classical use of the Viterbi algorithm for performing MLSD on a linear channel lies in the definition of the branch metrics (18), which incorporate the nonlinear features of the transmission channel in the reference branch labels $z'_{k\sigma+\eta}$. Also, reduced-state sequence detection (RSSD) algorithms [28]–[30] can be adapted to the proposed receiver structure, as will be done in Section V.

The number of trellis states grows exponentially with the channel memory L, where $(L+1)T$ is the effective support of the channel responses $f(t)$ and $\rho(t_1,t_2,t_3)$, with respect to each time variable. In a conventional linear system with baud-rate sampling ($\sigma=1$), a common approach is to design the filters so that the overall pulse $f(t)$ approximates a *bandlimited Nyquist pulse*, i.e., $f(nT)\simeq 0$ for $n\neq 0$. In this case, the time discrete equivalent channel may exhibit a drastically shorter memory, as compared to the effective duration of the time continuous overall pulse. On the other hand, the use of oversampling may result in a memory comparable with the effective length of this pulse. This effect must be regarded as an inherent feature of the oversampling front end and is not related to the presence of a nonlinear element. In fact, the channel memory would be the same if the oversampling strategy were applied in a linear channel with response $f(t)$. We note, however, that if the physical channel response $h(t)$ is strongly frequency selective, the effective duration of $f(t)$ is indeed a realistic measure of channel memory for any sampling rate.

IV. Error Performance Analysis

A performance analysis of MLSD receivers can be based on the classical approach described, for example, in [12] and [21]. Let an error event begin at time n, and assume that the MLSD receiver chooses a sequence $\hat{\mathbf{a}}$ when a different sequence $\mathbf{a}$ is transmitted. We define $\mathbf{e}\triangleq\hat{\mathbf{a}}-\mathbf{a}$ as the sequence of error symbols entailed by the error event. Hence, the path chosen by the receiver in the trellis diagram differs from the correct one by $D(\mathbf{e})$ state transitions, the first starting at time n, where $D(\mathbf{e})$ is called the duration of the error event, and the error sequence $\mathbf{e}=\{\ldots,0,0,e_n,e_{n+1},\ldots,e_{n+D(\mathbf{e})-L-1},0,0,\ldots\}$ includes $w(\mathbf{e})$ nonzero error symbols (at most, $D(\mathbf{e})-L$). The *square distance* of the error event is defined as

$$\begin{aligned}d^2(\hat{\mathbf{a}},\mathbf{a})&\triangleq\int_{T_0}|z'(\hat{\mathbf{a}},t)-z'(\mathbf{a},t)|^2dt\\&=\frac{T}{\sigma}\sum_{k=n}^{n+D(\mathbf{e})-1}\sum_{\eta=0}^{(\sigma-1)}|z'_{k\sigma+\eta}(\hat{\mathbf{a}})-z'_{k\sigma+\eta}(\mathbf{a})|^2\end{aligned} \tag{19}$$

where T_0 is the entire transmission period and the double summation extends only over the $\sigma D(\mathbf{e})$ nonzero terms. Note that the square distance is defined in terms of the time-continuous noiseless signals $z'(\mathbf{a},t)$ entering the sampler in Fig. 2, but it can also be computed in terms of the corresponding time-discrete Nyquist-rate-sampled sequences defined in (14).[2] Minimizing $d(\hat{\mathbf{a}},\mathbf{a})$ over all possible sequence pairs, $(\hat{\mathbf{a}},\mathbf{a})$ yields the minimum distance $d_{\min}$.

The following upper and lower bounds to the symbol error probability P_s hold in the general case of a deterministic channel affected by AWGN (e.g., see [21, eqs. (7.85) and (7.94)]):

$$P_s\leq\sum_{\mathbf{e}}w(\mathbf{e})\sum_{\mathbf{a}\in A(\mathbf{e})}Q\left[\frac{d(\mathbf{a}+\mathbf{e},\mathbf{a})}{\sqrt{2N_0}}\right]P\{\mathbf{a}\} \tag{20}$$

$$P_s\geq Q\left[\frac{d_{\min}}{\sqrt{2N_0}}\right]P\{\mathbf{a}\in A_{\min}\} \tag{21}$$

where $A(\mathbf{e})$ is the set of symbol sequences *compatible* with $\mathbf{e}$, i.e., those for which $\mathbf{a}+\mathbf{e}$ is an allowable symbol sequence, and $A_{\min}$ is the set of all symbol sequences minimizing the distance for some error sequence. Since at moderate and high signal-to-noise ratios (SNRs) the function $Q(\cdot)$ rapidly decreases with the distance, the upper bound (20) can be approximated by retaining only terms exhibiting a minimum distance

$$P_s\stackrel{\sim}{\leq}Q\left[\frac{d_{\min}}{\sqrt{2N_0}}\right]\sum_{\mathbf{e}\in E_{\min}}w(\mathbf{e})P\{\mathbf{a}\in A_{\min}(\mathbf{e})\} \tag{22}$$

where $E_{\min}$ and $A_{\min}(\mathbf{e})$ denote the sets of error sequences and corresponding information sequences with minimum distance, respectively. Comparing the lower bound (21) and the approximate upper bound (22), it appears that they differ for multiplicative constants only. Should these constants be equal or close, (22) would represent a good estimate of the error probability at high SNRs. In fact, these multiplicative constants are equal for the nonlinear channels considered in the numerical results.

In a classical approach to performance analysis for linear channels, (22) is considered a reliable approximation even at moderate values of SNR [12], [20], [22]. The minimization of $d^2(\hat{\mathbf{a}},\mathbf{a})$ with respect to both the estimated and transmitted sequence is a much more difficult task in a nonlinear channel than in the linear case; nonetheless, it can be performed by means of an algorithm, first proposed by Saxena [12], [23], that efficiently scans all possible path pairs in the trellis diagram, searching for

[2]Recall that a signal-nondistorting receiver filter $r(t)$ is assumed.

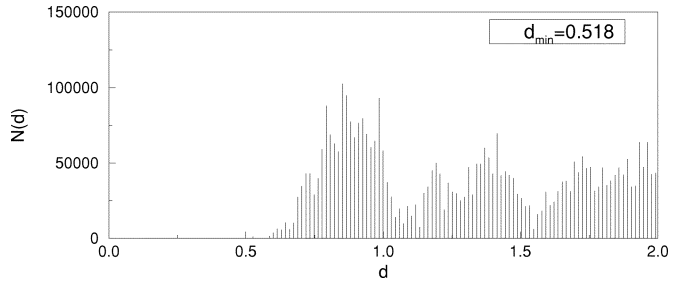

Fig. 3. Histogram of the distance spectrum for the nonlinear channel considered in Section V (d quantized in 100 contiguous intervals).

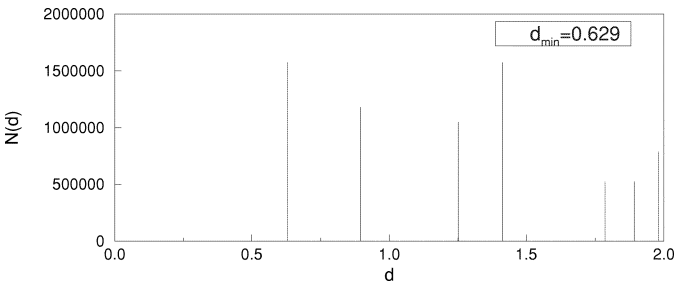

Fig. 4. Distance spectrum for the linearized channel considered in Section V.

the sequence pairs with the minimum distance $d_{\min}$. The reason for the increased complexity in this search is due to the lack of a uniform error property (UEP) that characterizes linear channels, for which the square distance only depends on the error sequence $\mathbf{e}$.

The absence of UEP in nonlinear channels has significant implications on the applicability of (22). It is better seen how (22) is a poor approximation by introducing the *distance spectrum* $N(d)$ of the nonlinear channel and comparing it with a similar spectrum of a linear channel obtained by perfect linearization. For a given sequence length and channel, $N(d)$ is defined as the number of unordered sequence pairs having distance d. Figs. 3 and 4 show examples of such spectra computed for the nonlinear channel considered at the beginning of Section V and its linearized counterpart, respectively. As is immediately recognized, for the same number of considered sequence pairs, the spectrum of the linearized channel is concentrated on some discrete values of the distance d, which correspond to the few error sequences most likely to occur. This is not the case for the nonlinear channel in Fig. 3, where for each value of the error sequence there is a variety of transmitted sequences $\mathbf{a}$ which yield different values for the distance $d(\mathbf{a}+\mathbf{e},\mathbf{a})$ and thus produce a *spreading effect* on the distance spectrum. The spreading is even more remarkable than what is visible in Fig. 3 since, for the sake of visualization, the figure is a histogram obtained by grouping the distance values in 100 contiguous intervals. In the linear channel of Fig. 4, it is possible to neglect all the spectrum lines, and the corresponding sequence pairs in (20), except the leftmost, with distance $d_{\min}$. On the contrary, it is not possible to do so in the nonlinear channel of Fig. 3, for which the distance spectrum shows a whole group of lines (the leftmost part of the spectrum) with distances d very close to $d_{\min}$, weighed by a multiplicity $N(d)$ much larger than $N(d_{\min})$ and giving significant contributions to (20).

It is also possible to rewrite the upper bound (20) by means of the distance spectrum. This is accomplished by partitioning all the $(\mathbf{a},\mathbf{e})$ pairs with respect to their distance. For every value of d appearing in the function $Q(\cdot)$ of (20), if we assume the same weight $w(d)$ for all error sequences associated with distance d, then $w(\cdot)$ and $Q(\cdot)$ can be factored out of the double summation. This assumption is verified on most channels of practical significance, such that the distance increases with the length of the error event and the relevant weight. The channels considered in the numerical results satisfy this assumption. For every value of d, the resulting double summation in (20) yields the probability of having $(\mathbf{a},\mathbf{e})$ pairs with this distance. This probability is trivially the ratio between the distance spectrum $N(d)$ and the total number N of unordered sequence pairs. For the sake of computability, one has to restrict the calculation of the spectrum to sequence pairs of a maximum given length. This is not a problem, since the distance increases with the error sequence length and large distance values can be neglected. This maximum length is the same length used for computing the total number N of sequence pairs. The new truncated upper bound then reads

$$P_s \overset{\sim}{\leq} \sum_{d=d_{\min}}^{d_{\max}} w(d)\frac{N(d)}{N}Q\left[\frac{d}{\sqrt{2N_0}}\right] \tag{23}$$

where the summation extends to a maximum value of the distance $d_{\max}$, above which the decreasing contributions to the summation can be neglected.

It is worth noting that the figures appearing on the vertical axes of Figs. 3 and 4 are justified as follows: starting from any of M^L trellis states, an error sequence of length $D(\mathbf{e}) = L+1$ and unit weight is assumed. One of the $M(M-1)$ error symbols is followed by L correctly detected symbols. With $M=16$ and $L=2$ as in Section V-A, the number of unordered sequence pairs of length $2L+1$ and unit weight is then $\left[M^L M(M-1)M^L\right]/2 \simeq 2^{23}$, equal to the sum of the distance spectrum discrete values.

In analogy with turbo codes [24], the multiplicity of pairwise errors with minimum distance [the only ones considered in the approximation (22)] is extremely small. These can be thought of as the only pairwise errors occurring when the SNR is extremely high and the symbol error rate (SER) is very low, where the two approximations in (22) and (23) tend to coincide. At the error rates of interest, error events with larger distance prevail, thus increasing the slope of the error curve at higher SERs. The dramatic truncation involved in (22) makes this approximation, and thus the significance of $d_{\min}$ for estimating the system performance, totally inadequate at the error rates of interest, as will be demonstrated in the numerical results of the Section V.

V. Numerical Results

In this section, we consider two different configurations for the transmission channel of Fig. 1, both employing a 16-QAM symbol set. The first includes an ideal passband physical channel filter $h(t)$ with 100% excess bandwidth; the second includes a seven-pole Chebyshev RF filter with cutoff frequency $1/2T$, equal to half the symbol rate. The latter choice accounts for the necessity to contain the spectrum of the nonlinearly amplified signal within an assigned bandwidth, thus avoiding

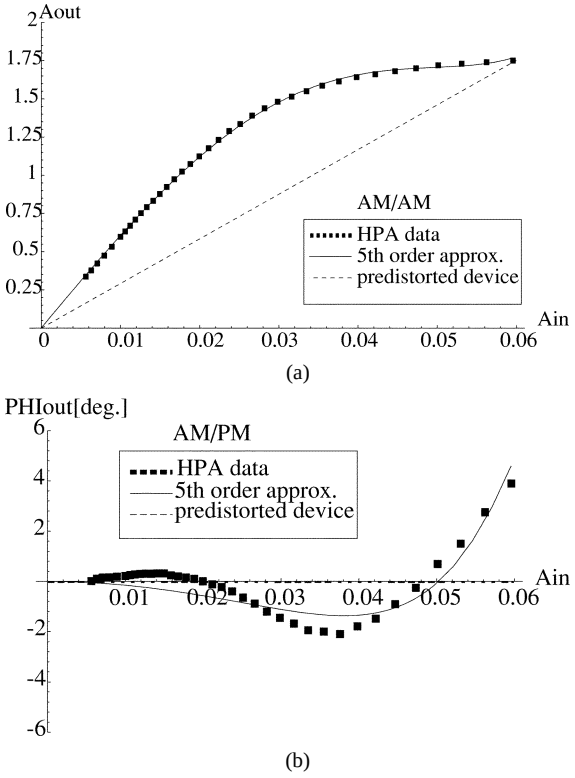

Fig. 5. AM/AM and AM/PM curves for the nonlinear semiconductor HPA considered in Section V-A and its fifth-order MMSE approximation; dashed lines show the ideally predistorted transfer characteristics.

interference with possible adjacent channels. This spectrum is not flat, in general, thus allowing the use of an RF filter with wider band. The bandwidth $1/2T$ is chosen as a *worst case* option. While the Chebyshev filter is physically realizable, its impulse response has a duration of several symbol periods and reduced complexity detection algorithms, such as RSSD, are needed in the receiver to limit the number of trellis states [28]–[30].

The simulation results on the SER will be compared, for each channel configuration, to the SER obtainable on a perfectly linearized channel including the same linear filters, in which the nonlinear block is replaced by a linear amplifier having an equal peak power level, i.e., with a linear AM/AM curve reaching the same peak level and an identically null AM/PM curve (see Fig. 5). This approach corresponds to an ideal analog predistortion of the nonlinear device, which could only be approximated in practice. The SER will be plotted versus the *peak-energy-to-noise ratio* E_p/N_0, where E_p is T times the square of the maximum value of the AM/AM curve of the amplifier.

The reason for comparing performance in terms of the peak energy E_p, instead of the average symbol or bit energy E_s or E_b, is that only the quantity E_p remains unchanged after the ideal analog predistortion. As for predistortion, the choice of fixing the peak transmitted power, instead of the average one, reflects a pragmatic point of view; in fact, it is not possible to linearize a device maintaining the same average output power (thus a larger peak power) without the aid of some supplemental, and possibly expensive, high-power device. With this choice, the nonlinear block inherently outputs a greater average power as compared to its predistorted counterpart with the same peak power. As a consequence, a comparison versus E_b/N_0 is of no practical interest, and the choice of E_p/N_0 is mandatory for a meaningful comparison. In Section V-A, we provide equations to relate E_b/N_0 and E_p/N_0 explicitly in the case of a linear channel.

A. Ideal Wideband Nonlinear Channel

Referring to Fig. 1, the nonlinear amplifier considered in this section will be based on data measured from a semiconductor high-power amplifier (HPA). In Fig. 5 the nonlinear transfer characteristics of this amplifier, labeled HPA, are shown using dots for measured data, together with the interpolating fifth-order polynomial nonlinear characteristics (solid line) having minimum mean square error (MMSE) from the measured data. Dashed lines show the transfer characteristics of a perfectly predistorted HPA, maintaining the same peak power, with linear AM/AM curve and identically null phase distortion (AM/PM).

The other elements of the transmission chain are a raised cosine shaping filter $p(t)$ with rolloff factor $\alpha = 0.7$ and a gain able to drive the NL element to full saturation; an ideal low-pass channel filter $h(t)$ without attenuation and one-sided bandwidth $1/T$ (100% excess bandwidth). Being $1/T$ the one-sided bandwidth of the noiseless received signal, an oversampling factor $\sigma = 2$ is sufficient if an ideal low-pass receiver front-end filter $R(f)$ with unit gain and one-sided bandwidth $1/T$ is chosen. Other choices with a smoother profile of $R(f)$ are possible, as outlined in Section III, but a larger σ results. The time-continuous response of the channel is, of course, infinite; for the purpose of receiver design, a truncation of the channel response is adopted, neglecting all linear and nonlinear channel dispersion parameters below 5% of their respective peak value [see (10) and (11)]. Based on this approximation, a time-discrete channel memory of $L = 2$ symbols results for both parameters, thus requiring a Viterbi processor with $M^L = 256$ trellis states. The truncation of channel memory implies neglecting some "tail effects" for each transmitted pulse.

Since the receiver is based on an approximate polynomial model of the amplifier, the receiver is not strictly optimal for the actual transmission channel. We also consider a reference system in which the real amplifier characteristic is replaced by its fifth-order approximation so that this receiver is optimal, in this case; this permits us to check the degradation margin involved in the polynomial approximation. In Fig. 6, the curve labeled HPA plots the SER against the peak-energy-to-noise power-spectral-density ratio E_p/N_0 for a system fed by the HPA semiconductor nonlinearity, with the receiver of Section III (designed on a fifth-order approximation of HPA). The curve labeled NL5 (nonlinearity of fifth-order) shows the performance of a similar system, where the power amplifier HPA is substituted by its fifth-order approximation (solid curves of Fig. 5) without altering the receiver, which is optimal in this case. The curve labeled L (linearized) identifies a system in which the HPA is perfectly predistorted, maintaining the same peak power (dashed curves of Fig. 5), and the receiver with oversampling of Section III operates in a *degenerate mode*, since there is not any nonlinear contribution to the received signal $z(t)$. Actually, since in this case $z'(t)$ is a linearly modulated 16-QAM signal with normalized signal set $(\mathrm{Re}[a_k] \in \{\pm 1; \pm 1/3\})$ and raised

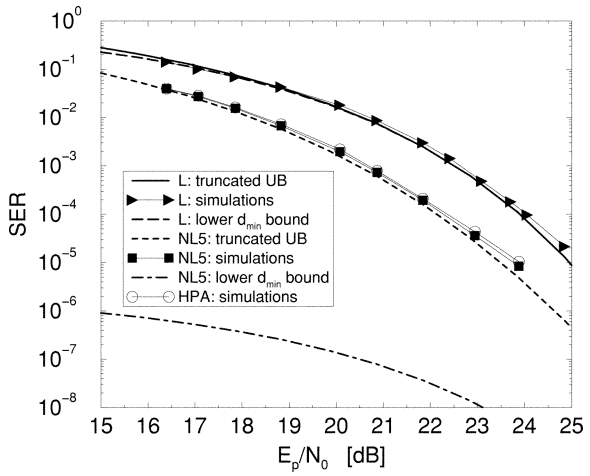

Fig. 6. Error performance and analytic approximations for the nonlinear channels considered in Section V-A (HPA and NL5) and their perfectly predistorted counterpart (L).

cosine pulse $f(t)$ with rolloff α and peak value $\sqrt{E_p/2T}$, performance is comparable with literature data, keeping in mind that

$$E_b = \frac{E_s}{4} = \frac{1}{4}E\{|a_k|^2\}\int |f(t)|^2 dt = \frac{1}{4}\frac{10}{9}\left(1-\frac{\alpha}{4}\right)\frac{E_p}{2} \tag{24}$$

as results from the integration of the squared raised cosine pulse; then, choosing a rolloff factor $\alpha = 0.7$, the following relation in logarithmic scale results:

$$E_{p,\mathrm{dB}} = E_{b,\mathrm{dB}} + 9.4. \tag{25}$$

For the linear (L) and fifth-order approximation (NL5) configurations in Fig. 6, the truncated upper bound (23) is also shown, in excellent agreement with simulation results; whereas the approximation in (21), labeled "lower $d_{\min}$ bound" in the figure, usually considered as a reliable estimate of SER [12], [25], is totally inadequate for these channels at the error rates of interest. The distance spectra of Figs. 3 and 4 have been used for calculating the two kinds of approximations. The significance of $d_{\min}$ for the linear case is demonstrated by the tightness of the truncated upper bound and the lower $d_{\min}$ bound even at high SERs. The better performance of the nonlinear channel can be intuitively justified by noting that even if the minimum distance of the nonlinear channel is lower, the *center of mass* of its distance spectrum is shifted to the right, with respect to the distance spectrum of the corresponding linearized channel.

From Fig. 6, the advantage of full utilization of the HPA dynamic range is evident. The loss entailed in employing a polynomial approximation of the HPA for the design of the receiver is negligible, as could have been conjectured from Fig. 5. A substantial gain of roughly 2 dB, for $\mathrm{SER} = 10^{-3}$, is achieved thanks to the greater average energy per transmitted symbol that the HPA is capable of delivering. Such an increased average energy is, on the other hand, responsible for the nonlinear distortion of the received pulses; the use of a suitable receiver is thus a key factor for efficiently exploiting the amplifier's potential.

The onset of a degradation at low SERs due to the "tail effects" induced by channel memory truncation can be inferred by the slight difference of the simulation curves with respect

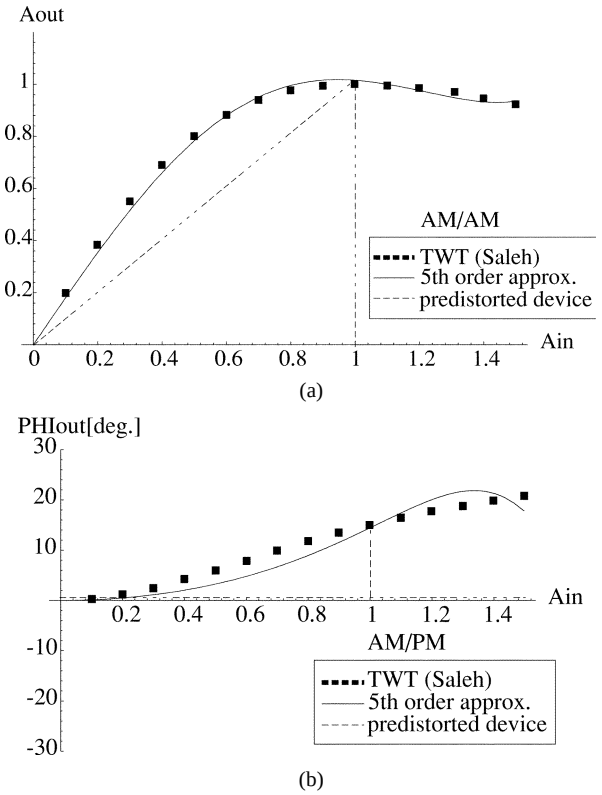

Fig. 7. (a) AM/AM and (b) AM/PM curves for the nonlinear TWT considered in Section V-A and its fifth-order MMSE approximation; dashed lines show the ideally predistorted transfer characteristics.

to theoretical predictions. Even simulations in the linear case suffer from such imperfections, since the same kind of receiver was employed for this purpose.

The asymptotical behavior of the two systems can be predicted using the following theoretical approximations. When noise approaches zero, the two truncated upper bounds (for linear and nonlinear channel) tend eventually to cross, and the ideally predistorted system performs better than the nonlinear one. Such a result is implicit in the comparison of the previously computed minimum distances, which become significant for the performance prediction of both systems only at very high E_p/N_0 ratios. This cross occurs at a SER of 10^{-11}, where simulations cannot be performed.

In order to demonstrate the robustness of the proposed receiver with respect to strong phase distortions, a traveling wave tube (TWT) amplifier is now considered in place of the semiconductor HPA employed so far. The Saleh model [26] is adopted, with an asymptotical phase shift of 30°, as shown in Fig. 7. The nonlinear characteristics of the TWT are nevertheless operated only up to the peak output power of the AM/AM curve, which is produced by an input amplitude $A_{\mathrm{in}} = 1$, and shows an AM/PM distortion of roughly 15°, as shown in Fig. 7, thus three times larger than the maximum phase distortion caused by the semiconductor HPA. Simulations show that TWT operation beyond this point does not improve performance.

Fig. 8 shows the simulation results for such a system and its perfectly predistorted counterpart. The curves labeled L coincide, of course, with those labeled L in Fig. 6, since the abscissa E_p/N_0 takes into account the remarkable difference in the peak

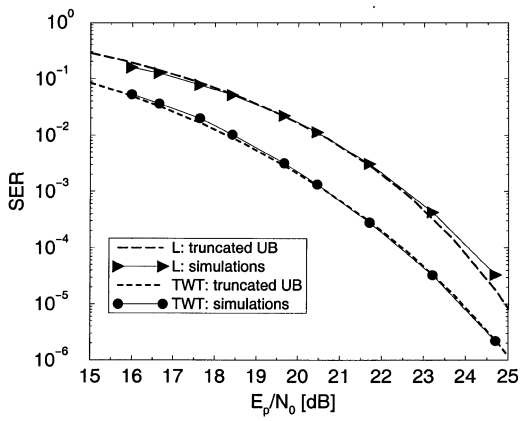

Fig. 8. Error performance and analytic approximations for the nonlinear channel considered in Section V-A (TWT) and its perfectly predistorted counterpart (L).

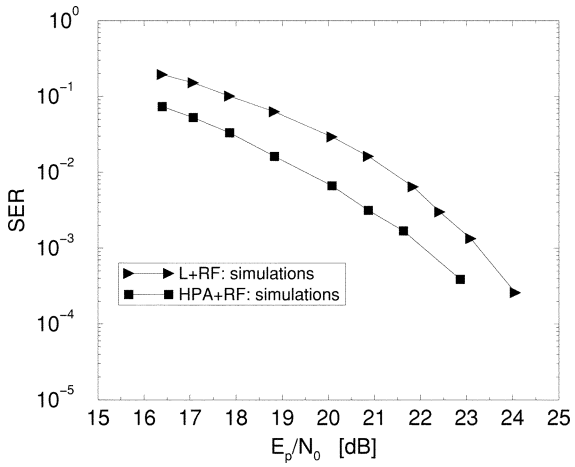

Fig. 9. Error performance for the nonlinear channel considered in Section V-B (semiconductor HPA plus a Chebyshev seven-pole RF filter) and its perfectly predistorted counterpart (L + RF).

output power of the two nonlinearities. It is noteworthy that the curve relative to the TWT system almost coincides with that of the HPA system of Fig. 6. This can be related to the relatively similar profile of the AM/AM curves for the two amplifiers, at least up to their peak point of utilization. This result suggests that the AM/AM curve plays a dominant role in the obtainable performance, whereas the AM/PM distortion is fully recovered by the proposed receiver. The simulation results are again in good agreement with the truncated upper bound approximations, computed on the respective distance spectra.

B. Narrowband Physical Channel

The main problem with the examined systems, which achieve a gain of roughly 2 dB as shown in Section V-A, is the requirement of a large physical bandwidth, namely double the symbol rate. In order to check whether a similar achievement is possible in the presence of strict limitations to the available bandwidth, an RF filter is inserted in the transmitter, cascaded to the previously described HPA amplifier.

The system under consideration consists of the same shaping filter, the same HPA and the same ideal physical channel filter of Section V-A, plus a seven-pole Chebyshev filter with passband *ripple amplitude* $A_r \simeq 0.26$ dB.[3] The (one-sided) cutoff frequency is first chosen as $0.8/T$ in order to avoid distortions on the linear signal component. The detection algorithm described in the following yields a loss of roughly 0.5 dB with respect to the ideal channel case of Section V-A, which has a cutoff frequency equal to $1/T$.

Since we wish to test the robustness of the receiver in the presence of severe bandwidth limitations, a *worst case* Chebyshev filter with cutoff frequency of $1/2T$ is chosen. This filter introduces distortion even on the linear part of the transmitted signal spectrum, and exhibits an attenuation of 40 dB at an offset frequency $0.7/T$ from the carrier, thus making the transmitted signal compliant even with *spectral emission masks* having strict requirements. The disadvantage of such a filter is that the total impulse response of the linear filters of this system, $f(t)$, has a much greater duration than in the wideband channel case. This is due to the passband ripples of the transfer function of the filter and, most of all, to the small cutoff frequency. If the effective time-continuous duration of $f(t)$ is measured with respect to a threshold equal to 5% of the peak value of $f(t)$, then a dispersion length $L = 6$ is obtained. To perform MLSD with full complexity with a 16-QAM modulation format, about 16 million states would be needed ($M^L = 2^{24}$). In this case, we can resort to RSSD techniques [28]–[30], typically employed when the problem of long channel memory is at hand.

A receiver similar to that described in Section III is employed, with $M^K = 256$ states, where $K = 2$ is the *reduced channel memory* considered in the definition of the trellis states. The noiseless samples $z'_{n\sigma+\eta}(a_{n-L}, \ldots, a_n)$ appearing in (18) are the sum of two separate contributions, the first being recovered from the state transition under consideration, which specifies only the K symbols $(a_{n-K}, \ldots, a_n)$, the second including the *cross contributions* that would appear in (14) when any of the summation indexes extends beyond K. The interfering symbols are in this case recovered from the trellis memory according to per-survivor processing [31]. The receiver is still based on a fifth-order approximation of the HPA. The cross contributions are then of three kinds (first-, third-, and fifth-order) and their number is extremely high. By considering only linear and nonlinear channel dispersion coefficients whose value is above 2% of their respective peak value, an affordable complexity is achieved while, again, introducing a truncation of the channel memory.

Fig. 9 shows the receiver performance for the considered narrowband nonlinear channel with semiconductor HPA. Again, performance is compared with a perfect analog predistortion approach, maintaining the same peak output power from the HPA. The gain margin in this worst case is reduced to 1 dB only, due to the severe approximations involved not only in the HPA polynomial representation, but also in channel memory truncation, for the sake of a simple receiver design. Gain margins obtainable by relaxing the strict worst-case condition on the RF filter bandwidth are, of course, larger.

The same RSSD approach has been applied to the wideband channel considered in Section V-A. Given a channel memory of

[3]The so called ϵ factor is equal to 0.25 [27].

$L = 2$, in that case, a simplified receiver considering a reduced channel memory of $K = 1$ symbol, and thus using a trellis with only 16 states, yields a loss of only 0.2 dB with respect to the curve labeled HPA, shown in Fig. 6.

VI. Conclusions

The design of a near optimal receiver for linearly modulated data sequences transmitted on nonlinear radio channels has been pursued in this paper. The use of a suitable front-end analog filter and signal oversampling renders the sequence of samples a sufficient statistic for performing MLSD, through the use of a VP. This solution is of practical interest, since only a modification of the constants appearing in the definition of the branch metrics is needed, without any modification in the structure of the processor.

The error performance of this MLSD receiver has been analyzed. An approximation to the symbol error probability through the use of the minimum distance has been shown to be inadequate for nonlinear channels. An enhanced approximation has been considered, based on the use of the full distance spectrum, showing that the lack of a uniform error property produces a spreading effect in the distance spectrum for nonlinear channels.

The performance of the proposed receiver has been compared, in terms of SER, with respect to that obtainable by employing perfect analog predistortion (which is in practice not feasible) of the nonlinear element, while maintaining its peak output power. For an ideal wideband physical channel, a gain of 2 dB has been achieved by the proposed MLSD receiver, both for a semiconductor HPA and a TWT, showing that the receiver is able to compensate for both phase and amplitude distortions. Simulations also demonstrated the adequacy of the proposed approximation of the error probability, as opposed to the little significance of the minimum distance approximation. An RF Chebyshev filter with *worst case* narrowband has been cascaded to the nonlinear amplifier. Using RSSD algorithms resulted in a lower, but still appreciable, gain margin.

References

[1] G. Karam and H. Sari, "A data predistortion technique with memory for QAM radio systems," *IEEE Trans. Commun.*, vol. 39, pp. 336–343, Feb. 1991.

[2] G. Lazzarin, S. Pupolin, and A. Sarti, "Nonlinearity compensation in digital radio systems," *IEEE Trans. Commun.*, vol. 42, pp. 988–999, Feb.-Apr. 1994.

[3] M.-G. Di Benedetto and P. Mandarini, "A new analog predistortion criterion with application to high efficiency digital radio links," *IEEE Trans. Commun.*, vol. 43, pp. 2966–2974, Dec. 1995.

[4] A. N. D'Andrea, V. Lottici, and R. Reggianini, "RF power amplifier linearization through amplitude and phase predistortion," *IEEE Trans. Commun.*, vol. 44, pp. 1477–1484, Nov. 1996.

[5] A. Bernardini and S. De Fina, "Analysis of different optimization criteria for IF predistortion in digital radio links with nonlinear amplifiers," *IEEE Trans. Commun.*, vol. 45, pp. 421–428, Apr. 1997.

[6] D. D. Falconer, "Adaptive equalization of channel nonlinearities in QAM data transmission systems," *Bell Syst. Tech. J.*, vol. 57, pp. 2589–2611, Sept. 1978.

[7] G. F. Hermann, "Performance of maximum-likelihood receiver in the nonlinear satellite channel," *IEEE Trans. Commun.*, vol. COM–26, pp. 373–378, Mar. 1978.

[8] K. Metzger and R. Valentin, "Intersymbol interference due to linear and nonlinear distortion," *IEEE Trans. Commun.*, vol. 44, pp. 809–816, July 1996.

[9] M. F. Mesiya, P. J. McLane, and L. L. Campbell, "Maximum-likelihood sequence estimation of binary sequences transmitted over bandlimited nonlinear channels," *IEEE Trans. Commun.*, vol. COM-25, pp. 633–643, July 1977.

[10] W. van Etten and F. van Vugt, "Maximum-likelihood receivers for data sequences transmitted over nonlinear channels," *Arch. Elektr. Übertragung.*, vol. 34, pp. 216–223, 1980.

[11] V. K. Dubey and D. P. Taylor, "Maximum-likelihood sequence detection for QPSK on nonlinear, band-limited channels," *IEEE Trans. Commun.*, vol. COM-34, pp. 1225–1235, Dec. 1986.

[12] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[13] A. Vannucci and R. Raheli, "Optimal sequence detection based on oversampling for bandlimited nonlinear channels," in *Proc. IEEE Int. Conf. Communication*, Atlanta, GA, May 1998, pp. 417–421.

[14] ——, "Nonlinear channels: Predistortion or enhanced detection?," in *Proc. IEEE Int. Conf. Universal Personal Communication*, Florence, Italy, Oct. 1998, pp. 1117–1122.

[15] D. R. Hummels and R. D. Gitchell, "Equivalent low-pass representation for bandpass Volterra systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 140–142, Jan. 1980.

[16] N. M. Blachman, "Detectors, bandpass nonlinearities, and their optimization: Inversion of the Chebyshev transform," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 398–404, July 1971.

[17] A. R. Kaye, D. A. George, and M. J. Eric, "Analysis and compensation of bandpass nonlinearities for communications," *IEEE Trans. Commun.*, vol. COM-20, pp. 965–972, Oct. 1972.

[18] H. Meyr, M. Oerder, and A. Polydoros, "On sampling rate, analog prefiltering, and sufficient statistics for digital receivers," *IEEE Trans. Commun.*, vol. 42, pp. 3208–3214, Dec. 1994.

[19] P. Castoldi and R. Raheli, "On recursive optimal detection of linear modulations in the presence of random fading," *Eur. Trans. Telecommun. (ETT)*, vol. 9, pp. 209–220, Mar./Apr. 1998.

[20] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.

[21] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Norwell, MA: Kluwer, 1988.

[22] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun.*, vol. COM-22, pp. 624–636, May 1974.

[23] M. G. Mulligan and S. G. Wilson, "An improved algorithm for evaluating trellis phase codes," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 846–851, Nov. 1984.

[24] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 409–428, Mar. 1996.

[25] A. Gutierrez and W. E. Ryan, "The performance of MLSD receivers on nonlinear satellite channels," in *Proc. Int. Conf. Communications*, 1996, pp. 921–925.

[26] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. COM–29, pp. 1715–1720, Nov. 1981.

[27] A. B. Williams and F. J. Taylor, *Electronic Filter Design Handbook*. New York: McGraw-Hill, 1988.

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

[29] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.

[30] P. R. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. Commun.*, vol. 37, pp. 669–676, July 1989.

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

Armando Vannucci (S'95–M'01) was born in Frosinone, Italy, in 1968. He received the Dr. Ing. (*Laurea*) degree in electronics engineering (*summa cum laude*) from the University of Roma "La Sapienza," Rome, Italy in 1993, and the Ph.D. degree in information engineering from the University of Parma, Parma, Italy in 1998.

Until 1995 he was with the INFO-COM Department, University of Roma, Rome, Italy, doing research activity in the field of acoustic phonetics. Since 1995, he has been with the Dipartimento di Ingegneria dell'Informazione, University of Parma, conducting research in the field of nonlinear radio channels from 1995 to 1998. Since 1999, his research interests have been in the field of optical transmission and optical communication systems.

Riccardo Raheli (M'87) received the Dr. Ing. degree (Laurea) in electrical engineering (*summa cum laude*) from the University of Pisa, Pisa, Italy in 1983, the M.S. degree in electrical and computer engineering from the University of Massachusetts at Amherst in 1986, and the Ph.D. degree (Perfezionamento) in electrical engineering (*summa cum laude*) from the Scuola Superiore di Studi Universitari e di Perfezionamento (now "S. Anna"), Pisa, Italy in 1987.

From 1986 to 1988, he was a Project Engineer at Siemens Telecomunicazioni, Cassina de' Pecchi, Milan, Italy. From 1988 to 1991, he was a Research Professor at the Scuola Superiore di Studi Universitari e di Perfezionamento S. Anna. In 1990, he was a Visiting Assistant Professor at the University of Southern California, Los Angeles. Since 1991, he has been with the University of Parma, Parma, Italy, where he is currently a Professor of Telecommunications. His scientific interests are in the general area of statistical communication theory, with special attention on digital transmission systems, data sequence detection techniques, digital signal processing, and adaptive algorithms for telecommunications. His research activity has lead to numerous scientific publications in leading international journals and conference proceedings and a few industrial patents.

Dr. Raheli has served on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS as an Editor for Detection, Equalization, and Coding since 1999.