Optimal Sequence Detection Based on Oversampling for Bandlimited Nonlinear Channels

Armando Vannucci and Riccardo Raheli
Dipartimento di Ingegneria dell'Informazione, Università di Parma, I - 43100 Parma, Italy

Abstract— **Based on a polynomial representation of a memoryless bandpass nonlinearity, a new realization of an optimal receiver is proposed to perform maximum likelihood detection of data sequences transmitted over nonlinear, possibly time-dispersive, channels. The receiver employs oversampling of the observed signal to compute proper branch metrics for a Viterbi processor. Error performance is compared to that of an optimal receiver for the linear channel obtained by ideal analog predistortion of the nonlinear device under a peak-power constraint. In the presence of nonlinear distortion, a significant improvement in the symbol error rate is shown to be achievable by optimal detection with respect to ideal predistortion. The numerical results are based on both analytic and simulation methods.**

I. Introduction

THE performance of bandpass digital communication systems in the presence of nonlinear distortion due to high power amplifiers (HPA), such as travelling-wave tubes or solid-state devices, has traditionally been a controversial topic. In several papers, claims have been made that the error performance of these systems is not inherently degraded by the nonlinear distortion [1], or even that there may be a substantial gain especially when the channel is time-dispersive [2].

The main problem with the considered channels is the design of an appropriate receiver of affordable complexity able to recover the joint effects of nonlinear distorsion and intersymbol interference (ISI) due to both the HPA and the possibly limited physical channel bandwidth. Several receiver structures have been proposed [3]-[5] for performing maximum likelihood sequence detection (MLSD) in the presence of bandlimited nonlinear channels; one problem with all the proposed structures is the dramatic increase in the number of matched filters needed in the receiver front-end as the dispersion length of the channel and/or the order of the nonlinearity grows.

The aim of this work is to propose a new optimal MLSD receiver structure for linearly modulated signals which uses only one filter in the receiver front-end, followed by a sampler with rate larger than the signaling frequency by a so-called "oversampling" factor. Under proper conditions on the frequency response of the receiver filter and the oversampling factor, a sufficient statistics for data detection is obtained [6]. For a specific possible choice, an uncorrelated sampled noise sequence may be obtained, with a significant complexity saving in the receiver. The sampled sequence of observations is optimally processed by a Viterbi algorithm with proper branch metrics. The main complexity resides in the branch metric unit, which needs to have appropriate knowledge of both channel response and nonlinearity.

The error performance of an optimal receiver is difficult to estimate analytically, since the pairwise error probability (PEP) between sequences does not depend on the error sequence only, but also on the particular transmitted sequence. Approximate upper and lower bounds to the symbol error probability (SEP) are proposed. It is shown that the minimum Euclidean distance between sequences does not represent a realistic performance estimate. Simulation results based on the proposed receiver structure are provided to complement the analytical findings in a simple ISI channel with a dispersion length of two symbol intervals and a 16-QAM (quadrature amplitude modulation) signal set. The accuracy of the derived bounds is shown. The numerical results clearly show that a significant gain can be achieved by optimal sequence detection with respect to ideal predistortion. The proposed receiver appears also in [7].

II. System model

Let us consider the baseband equivalent of a bandpass digital communication system, shown in fig. 1, which transmits a sequence $\{a_k\}$ of independent and uniformly distributed symbols belonging to a discrete complex M-ary alphabet. The shaping filter with impulse response $p(t)$ is followed by a bandpass nonlinearity (NL) that feeds the distorted signal $y(t)$ to a linear system with impulse response $h(t)$. The latter filter can be viewed as the cascade of a radio frequency (RF) filter, which may limit the spectral occupation of the transmitted signal, and the physical channel. The received signal $s(t)$ is assumed to be affected by additive white Gaussian noise (AWGN) $n(t)$. Our primary problem is to find a suitable receiver structure (REC in the figure) that yields a maximum likelihood estimate $\{\hat{a}_k\}$ of the transmitted sequence.

In order to obtain a simple analytic representation of the received signal $s(t)$, we need a model for the baseband equivalent of the nonlinear memoryless device. Kaye et al. [8] showed that a *quadrature model*, including different memoryless nonlinearities in two parallel in phase and quadrature branches, can account for both amplitude distortion (AM/AM) and phase distortion (AM/PM) usually experienced at the output of bandpass nonlinear devices—both effects cannot be otherwise analytically described by a single nonlinear function. A polynomial approximation of the nonlinearities in both branches seems appropriate since it can be shown [9] that devices with such transfer characteristics maintain the same polynomial structure when rep-

This work was performed partly within a research cooperation between the Dipartimento di Ingegneria dell'Informazione, Università di Parma, Italy and Italtel S.p.A., Milano, Italy.

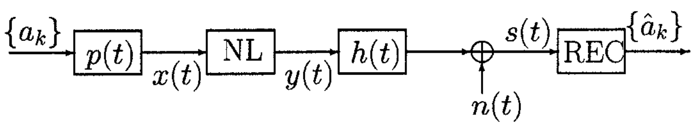

Fig. 1. Baseband equivalent of a nonlinear system with AWGN.

resented in terms of input and output complex envelopes. Moreover, the baseband equivalent of the quadrature model can be described as a single polynomial function with complex coefficients γ_i [10]. Hence, the relationship between the complex envelopes $x(t)$ and $y(t)$ reads

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

in which only odd powers of $x(t)$ are present due to the bandpass nature of the nonlinearity [9].

The linearly modulated signal $x(t)$ with signaling period T is expressed as

$$x(t) = \sum_n a_n p(t - nT) \quad . \qquad (2)$$

Substituting (2) in (1) and assuming a third-order nonlinearity ($N = 1$) for simplicity, signal $y(t)$ may be expressed as

$$y(t) = \gamma_1 \sum_n a_n p(t - nT) + \frac{3}{4}\gamma_3 \sum_i \sum_j \sum_l a_i a_j a_l^* \cdot p(t - iT) p(t - jT) p^*(t - lT) \quad . \qquad (3)$$

This expression may be generalized to account for higher order terms in a straightforward manner. In the following, we only consider third-order nonlinearities as it is sufficient to outline all the relevant concepts.

Taking into account the filtering effect of $h(t)$, we can express the received signal $s(t)$ in the compact form

$$s(t) = \gamma_1 \sum_n a_n f(t - nT) + \frac{3}{4}\gamma_3 \sum_i \sum_j \sum_l a_i a_j a_l^* \cdot \rho(t - iT, t - jT, t - lT) + n(t) \qquad (4)$$

having defined

$$f(t) \triangleq \int_{-\infty}^{+\infty} h(\tau) p(t - \tau)\, d\tau \qquad (5)$$

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

which may be regarded as a first-order and third-order impulse response of the nonlinear channel. It may be noted that our assumptions are equivalent to assuming a Volterra model for the cascade of a memoryless nonlinearity and linear filters: the corresponding Volterra kernels are simply found by multiplying the coefficients γ_i and the functions $f(t)$, $\rho(t_1, t_2, t_3)$.

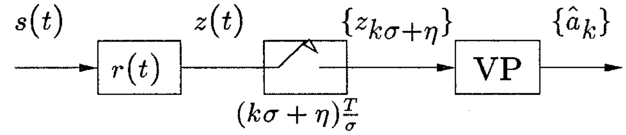

Fig. 2. Optimal receiver structure.

III. Optimal receiver structure

Let us assume the receiver structure shown in fig. 2. The received signal is passed through a receiver front-end filter with impulse response $r(t)$. The filtered signal $z(t)$ is sampled at a rate $\frac{\sigma}{T}$, multiple of the signaling frequency by an appropriate integer σ. Under proper conditions on $r(t)$ and σ, the sampled sequence is a sufficient statistics for the estimation of the information sequence. These conditions require that the transfer function $R(f)$ of the receiver filter is nonzero over the entire bandwidth of $s(t)$ and that it is strictly bandlimited [6]. We note that the bandwidth of $s(t)$ might be larger than the bandwidth of the modulated signal $x(t)$, due to the nonlinearity, but smaller than the bandwidth of the distorted signal $y(t)$, due to the RF and channel filtering $h(t)$. A sufficient condition is that $R(f)$ is constant in the signal bandwidth and $|R(f)|^2$ satisfies the first Nyquist criterion for the absence of ISI for a signaling frequency $\frac{\sigma}{T}$. In this case, the observed samples are also affected by independent Gaussian noise samples.

The observed signal $z(t)$ is thus oversampled, with respect to the signaling rate, to obtain a set of samples $z_{k\sigma+\eta}$ whose subscript is conveniently written using two integer symbols: k identifies the transmitted symbol period and η scans the σ samples obtained in the k-th period, η taking values in the integer set $\{0; \sigma - 1\}$. Under these conditions, it is always possible to reconstruct the noiseless component of the received signal $s(t)$ from the observed samples; hence, they constitute a sufficient statistics for the estimation of the information sequence. The described receiver filter is one of the possibilities suggested by the general approach described in [6] to obtain a sufficient statistics, which gives advantages in receiver implementation since it yields uncorrelated noise samples.

By proper algebraic manipulation, the observed samples can be expressed in the following form

$$z_{k\sigma+\eta}(a_{k-L} \cdots a_k) = \gamma_1 \sum_{n=0}^{L} f_{n\sigma+\eta}^{(\sigma)} a_{k-n} + \frac{3}{4}\gamma_3 \cdot \sum_{i=0}^{L} \sum_{j=0}^{L} \sum_{l=0}^{L} \rho_{i\sigma+\eta, j\sigma+\eta, l\sigma+\eta}^{(\sigma)} a_{k-i} a_{k-j} a_{k-l}^* + w_{k\sigma+\eta} \qquad (7)$$

where we have defined the following parameters

$$f_{n\sigma+\eta}^{(\sigma)} \triangleq f(nT + \eta\frac{T}{\sigma}) \qquad (8)$$

$$\rho_{i\sigma+\eta, j\sigma+\eta, l\sigma+\eta}^{(\sigma)} \triangleq \rho(iT + \eta\frac{T}{\sigma}, jT + \eta\frac{T}{\sigma}, lT + \eta\frac{T}{\sigma}) \qquad (9)$$

which can be regarded as oversampled first-order and third-order channel dispersion parameters. Their value can be computed through the expressions (5) and (6). These equations do not need any modification to account for the receiver filter because of the assumption that its frequency response is flat in the bandwidth of interest; if this is not the case, the function $h(t)$ should be modified to include the cascade of both channel and receiver filters and the receiver should take into account this distortion in the subsequent processing. Parameter L in (7) is defined as the number of symbol periods in which the value of the first and third order time-continuous impulse responses, here assumed causal, is not negligible; thus, it defines the dispersion length of the equivalent time-discrete nonlinear channel. Under the assumption of system filters with strictly finite duration, L is exactly the duration of the linear channel impulse response which would result if the nonlinear device was removed or perfectly predistorted. In fact, if the integrand in (5) is zero for $t > LT$, also the integrand in (6) is zero for any $t_i > LT$. If the physical channel is ideal or moderately frequency-selective, the actual time-discrete dispersion length of the linearized channel may be less than L, but under the assumption of severe frequency selectivity, the dispersion length of the first-order and third-order time-discrete channel parameters (8) and (9) may actually coincide. The above statements have been numerically verified for various channel and transmitter configurations and the summation limits in (7) are thus justified.

Since the quantities $w_{k\sigma+\eta}$ are uncorrelated Gaussian noise samples, the observations $\{z_{k\sigma+\eta}\}$, denoted by a vector $\mathbf{z}$, have a multivariate Gaussian distribution with mean value $\mathbf{z}'$ and uncorrelated elements, all with the same variance σ_w^2 corresponding to the variance of the noise samples. The elements $z'_{k\sigma+\eta}$ of the mean vector $\mathbf{z}'$ trivially correspond to the noiseless part of the observed samples (7). Estimation of the transmitted sequence under the ML criterion requires finding the sequence $\hat{\mathbf{a}} \triangleq \{\hat{a}_k\}$, among all the possible transmitted sequences $\mathbf{a} \triangleq \{a_k\}$, which maximizes the probability density function $p(\mathbf{z}|\mathbf{a})$ of the observed vector $\mathbf{z}$ conditioned on the transmission of the particular sequence $\mathbf{a}$. The maximization of the Gaussian probability density function $p(\mathbf{z}|\mathbf{a})$ is equivalent to the minimization of the absolute value of its exponent. Thus, if a finite-length transmitted sequence extends over $N+1$ symbols, the MLSE decision rule becomes

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}} \left\{ \sum_{k=0}^{N+L} \sum_{\eta=0}^{\sigma-1} |z_{k\sigma+\eta} - z'_{k\sigma+\eta}(\mathbf{a})|^2 \right\} \tag{10}$$

Since the noiseless samples $z'_{k\sigma+\eta}(a_{n-L}\cdots a_n)$ only depend on a finite sequence of symbols (see (7)), it is possible to express the sequence metric (10) to be minimized in a recursive form. We define the partial sequence metric

$$\begin{aligned}\Lambda_{n+1}(\mathbf{a}) &\triangleq \sum_{k=0}^{n} \sum_{\eta=0}^{\sigma-1} |z_{k\sigma+\eta} - z'_{k\sigma+\eta}(\mathbf{a})|^2 \\ &= \Lambda_n(\mathbf{a}) + \lambda(\mu_n \to \mu_{n+1})\end{aligned} \tag{11}$$

recursively updatable through the incremental branch metrics

$$\lambda(\mu_n \to \mu_{n+1}) \triangleq \sum_{\eta=0}^{\sigma-1} |z_{n\sigma+\eta} - z'_{n\sigma+\eta}(a_{n-L}\cdots a_n)|^2 \tag{12}$$

defined on a classical trellis diagram whose states $\mu_n \triangleq (a_{n-L}\cdots a_{n-1})$ coincide with the content of the discrete channel memory. The trellis branches, or similarly the state transitions $(\mu_n \to \mu_{n+1})$, specify the set of symbols $(a_{n-L}\cdots a_n)$ and the σ reference samples $z'_{n\sigma+\eta}$ at the n-th symbol period. The minimization in (10) can thus be conveniently performed by means of the well-known Viterbi algorithm with M^L states implemented through a Viterbi processor (VP in fig. 2) whose branch metric unit computes the quantities in (12).

IV. Performance bounds for an optimal receiver

In order to determine the performance of the proposed receiver, we can resort to the classical concept of error event: an error event begins at discrete time n whenever the receiver detects a sequence $\hat{\mathbf{a}} = \mathbf{a}+\mathbf{e}$ different from the actual transmitted sequence $\mathbf{a}$ and the error sequence $\mathbf{e}$ entails a finite number $D(\mathbf{e}) - L$ of consecutive nonzero symbols, the first being in the n-th position; this causes $D(\mathbf{e})$ wrong transitions in the trellis. It can be shown that, for the proposed receiver, or in general an optimal one, the probability that the path metric of the wrong sequence $\mathbf{a}+\mathbf{e}$ is less than that of the actual sequence $\mathbf{a}$ can be expressed as

$$P\{\Lambda(\hat{\mathbf{a}}) \le \Lambda(\mathbf{a})\} = Q\left[\frac{d(\hat{\mathbf{a}},\mathbf{a})}{\sqrt{2N_0}}\right] \tag{13}$$

where $\Lambda(\mathbf{a})$ denotes the metric of sequence $\mathbf{a}$, $\frac{N_0}{2}$ is the power spectral density of the real (or imaginary) part of the noise $n(t)$, assumed complex Gaussian, and we have defined the following Euclidean distance in the signal space between observed sequences

$$\begin{aligned}d^2(\hat{\mathbf{a}},\mathbf{a}) &\triangleq \frac{T}{\sigma} \sum_{k=n}^{n+D(\mathbf{e})-1} \sum_{\eta=0}^{(\sigma-1)} |z'_{k\sigma+\eta}(\hat{\mathbf{a}}) - z'_{k\sigma+\eta}(\mathbf{a})|^2 \\ &= \int_{T_0} |z'(\hat{\mathbf{a}},t) - z'(\mathbf{a},t)|^2 dt \quad ;\end{aligned} \tag{14}$$

the dependence on the transmitted sequence is shown and the integral extends over the entire transmission period T_0. The first equality in (14) is directly derived from the decision rule of the described receiver, whereas the second equality states that the distance can be viewed also in the time-continuous domain as the energy of the difference between the noiseless signals conveying the two sequences. The latter expression obviously coincides with the usual definition of signal distance found in the existing literature (e.g. [5]). The second equality in (14) holds since the signals $z'(\mathbf{a},t)$ are sampled at their Nyquist rate. Despite the possibly large bandwidth of the receiver filter, it is only N_0 which influences the probability in (13) and not the variance σ_w^2.

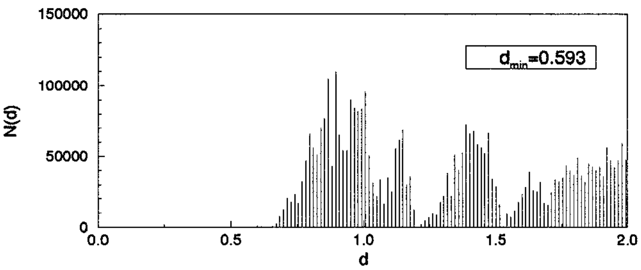

Fig. 3. Distance spectrum for the nonlinear channel considered in section V.

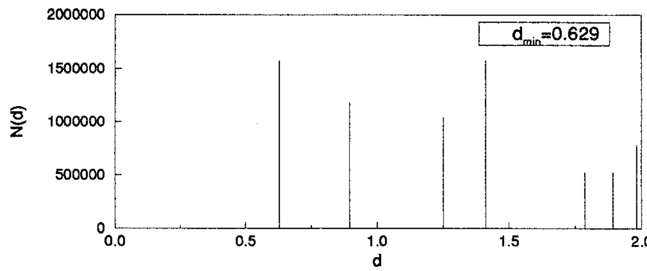

Fig. 4. Distance spectrum for the linearized channel considered in section V.

The pairwise error probability (PEP) $P(\mathbf{a}+\mathbf{e}|\mathbf{a})$ of estimating $\mathbf{a}+\mathbf{e}$ when $\mathbf{a}$ is transmitted can be upper bounded by (13). Thus we can express an error event E through a union over all the possible error sequences $\mathbf{e}$ starting at time n and then derive the following union bound

$$P\{E\} \leq \sum_{\mathbf{e}} \sum_{\mathbf{a}\in A(\mathbf{e})} Q\left[\frac{d(\mathbf{a}+\mathbf{e},\mathbf{a})}{\sqrt{2N_0}}\right] P\{\mathbf{a}\} \qquad (15)$$

where $A(\mathbf{e})$ is the set of all possible transmitted sequences compatible with $\mathbf{e}$, i.e., $\mathbf{a} \in A(\mathbf{e})$ if all the symbols of $\mathbf{a}+\mathbf{e}$ belong to the signal set.

Unfortunately, in nonlinear channels the uniform error property does not hold since the distance (14) does not depend on the error sequence only, as it does for linear channels. We could now search for the minimum distance d_{min} between sequences by minimizing (14) with respect to both sequences $(\hat{\mathbf{a}}, \mathbf{a})$: this can be done through an algorithm first used by Saxena [5], [11]. The resulting truncated upper bound is

$$P\{E\} \simeq Q\left[\frac{d_{min}}{\sqrt{2N_0}}\right] \sum_{\mathbf{e}\in U(d_{min})} P\{\mathbf{a} \in A(\mathbf{e})\} \qquad (16)$$

in which $U(d_{min})$ is the set of error sequences which minimize the distance for some $\mathbf{a}$. Unfortunately, this bound does not account for realistic performance since there may be a large number of other sequence couples giving rise to pairwise errors having distance slightly larger than d_{min}; these couples cannot be neglected, in general, and the right hand side of (16) is only an asymptotic lower bound of $P\{E\}$. The probability $P\{\mathbf{a} \in A(\mathbf{e})\}$ is trivially the ratio between the number of sequences compatible with $\mathbf{e}$ and the total number of sequences; the search can be obviously confined to sequences of length equal to the duration $D(\mathbf{e})$ of the error event.

A *distance spectrum* $N(d)$ is defined as the number of sequence pairs with distance d as a function of d. For a linear channel, there are only a few distance values, corresponding to respective error sequences; among these, the one associated with d_{min} dominates the probability $P\{E\}$ since the distance values differ significantly. As shown in the next section, for the nonlinear channel there is a sort of *spreading effect* which generates a large number of distance values for a given error sequence and different transmitted sequences. The distance spectrum $N(d)$ can be used as a weight function by rewriting the truncated upper bound (15) as a summation over the distance d extending from d_{min} to some value d_{max} for which the contributions of the terms $Q[\cdot]$ become negligible:

$$P\{E\} \simeq \sum_{d=d_{min}}^{d_{max}} \frac{N(d)}{N} Q\left[\frac{d}{\sqrt{2N_0}}\right] \qquad (17)$$

where N is the total number of unordered sequence pairs of given length.

V. Numerical results

For the purpose of simulation, a simple nonlinear channel with the structure of fig. 1 has been considered. The transmitter filter has a raised-cosine transfer function $P(f)$ with roll-off equal to 0.7. A third-order nonlinearity is assumed, with amplitude and phase distortion as shown in fig. 5. Both $h(t)$ and $r(t)$ are ideal bandpass filters with bandwidth $\frac{2}{T}$, and consequently an oversampling factor $\sigma = 2$ is employed. The information symbols belong to a 16-QAM constellation. These system choices led to a dispersion length L of 2 symbol periods with a resulting trellis diagram with 256 states.

Figures 3 and 4 show the distance spectrum for the described nonlinear channel and the linear channel obtained, from the latter, by perfect predistortion of the NL device, respectively. It is evident that the distance spectrum for the nonlinear channel undergoes a spreading effect, as mentioned in the last section.

Since it has been noted by inspection that in this channel the error sequence pairs with distance close to d_{min} always involve a single nonzero error symbol, the probability $P\{E\}$ of an error event is assumed to coincide with the symbol error rate (SER). The error events dominating the SER have a duration $D(\mathbf{e}) = 3$.

Fig. 6 shows the simulated SER versus the inverse of N_0 along with the previously introduced truncated upper bound (UB) (17) and asymptotical lower bound (16). From a pragmatic point of view, the signal to noise ratio $\frac{E_b}{N_0}$ is not a significant value for our comparison since a predistortion of the nonlinearity that maintains the same energy per transmitted bit is impossible without the aid of some supplemental high-power device. The abscissa $\frac{1}{N_0}$ used in fig. 6 is proportional to the peak-power to noise-power ratio.

From the figure, it is clearly seen that the approximation in (16), by means of d_{min} only, is of little significance at the error rates of interest. The truncated UB was derived with the following approximations: only error events with length

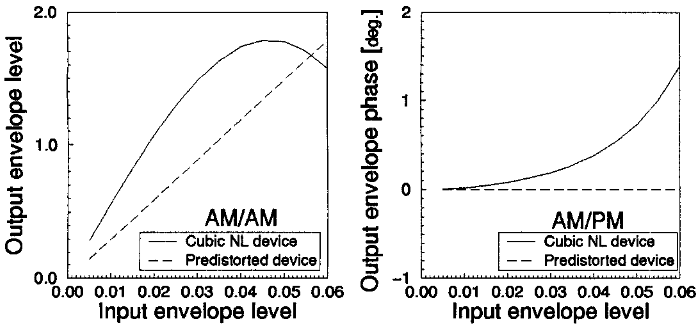

Fig. 5. AM/AM and AM/PM for the cubic nonlinearity considered in the simulation; dashed lines show the ideally predistorted transfer characteristics.

3 and distance in the range $[d_{min}; d_{max}]$ were considered; the filters considered for the calculation of distances are of strictly finite duration, thus disregarding the "tail effects" of the actual finite bandwidth filters employed in the channel.

The same analysis has been performed for a linear channel obtained by a perfect analog predistorsion of the nonlinearity up to the maximum peak-power level of the input-output characteristic, as shown in fig. 5. The truncated UB and the lower bound are hardly distinguishable in this case, because of the significance of the minimum distance parameter for linear channels.

A gain margin of roughly 2.5 dB is shown at a SER of 10^{-3}; this gain eventually reduces at lower error rates but is still greater than zero for SERs greater than 10^{-30}, where the truncated UBs for the two considered channels eventually cross. It is implicit in the assumption of equal peak-power that the average transmitted energy per bit is greater in the nonlinear channel than in the predistorted one. In fact, this is the reason for the gain shown.

VI. Conclusions

In this paper a new optimal receiver structure, based on oversampling, which performs MLSD over nonlinear time-dispersive channels has been introduced. It includes a Viterbi processor employing branch metrics (12). If the physical channel entails significant ISI effects, the state-complexity of the VP is not increased by the presence of the nonlinear device.

A truncated upper bound and an asymptotical lower bound have been derived for an optimal receiver. The little practical significance of the minimum Euclidean distance between sequences for a nonlinear channel has been demonstrated. In order to derive tighter approximations of the actual symbol error rate the use of a *distance spectrum* is required.

Simulations have been performed to compare the error performance of the proposed receiver to that of an optimal MLSD receiver for the perfectly linearized channel under a peak-power constraint. The accuracy of the theoretical analysis has been verified. A gain margin of more than 2 dB has been shown for the proposed receiver with respect to an optimal receiver employed on the predistorted channel, for values of SER of practical interest.

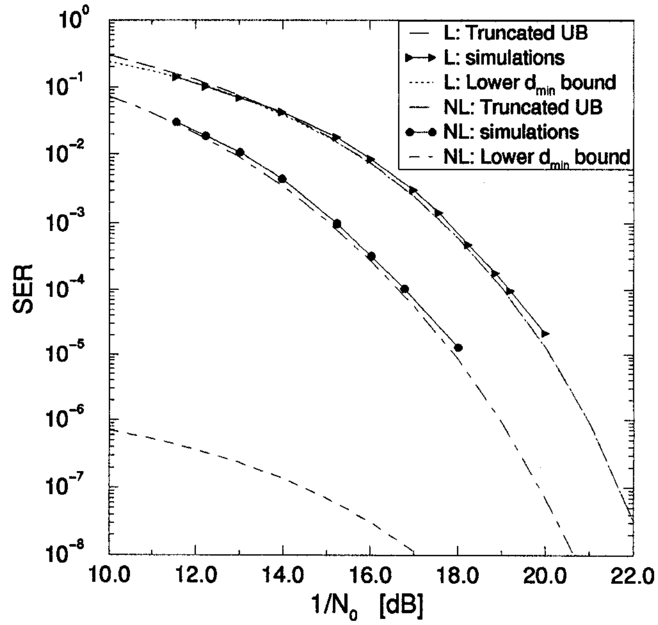

Fig. 6. Error performance and analytic approximations for cubic (NL) and predistorted (L) channel.

The results show that it is not always convenient to adopt analog predistortion of high-power nonlinear devices but rather to exploit fully the available power through the use of an enhanced receiver which takes into account the nonlinear effects.

References

[1] G. F. Hermann, "Performance of Maximum-Likelihood Receiver in the Nonlinear Satellite Channel", *IEEE Trans. Commun.*, vol. COM-26, pp.373-378, March 1978.

[2] K. Metzger, R. Valentin, "Intersymbol Interference Due to Linear and Nonlinear Distortion", *IEEE Trans. Commun.*, vol. COM-44, pp.809-816, July 1996.

[3] 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.

[4] W. van Etten and F. van Vugt, "Maximum Likelihood Receivers for Data Sequences Transmitted over Nonlinear Channels", *Archiv für Elektronik und Übertragungsteknik (AEÜ)*, vol.34, pp.216-223, 1980.

[5] S. Benedetto, E. Biglieri, V. Castellani, *Digital Transmission Theory*, Prentice-Hall, 1987.

[6] H. Meyr, M. Oerder and A. Polydoros, "On Sampling Rate, Analog Prefiltering, and Sufficient Statistics for Digital Receivers", *IEEE Trans. Commun.*, vol. COM-42, pp.3208-3214, Dec. 1994.

[7] R. Raheli and A. Vannucci, Ricevitore a stima di sequenza a massima verosimiglianza per la ricezione di dati seriali trasmessi su canali di comunicazione comprendenti una non linearità polinomiale, Italtel S.p.A., Milano, Italy, Italian patent n. MI97A 001299, June 1997.

[8] A. R. Kaye, D. A. George and M. J. Eric, "Analysis and Compensation of Bandpass Nonlinearities for Communications", *IEEE Trans. Commun.*, vol. COM-30, pp.965-972, October 1972.

[9] 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.

[10] S. Benedetto, E. Biglieri, R. Daffara, "Modeling and Performance Evaluation of Nonlinear Satellite Links - A Volterra Series Approach", *IEEE Trans. Aerospace Elec. Systems*, vol. AES-15, pp.494-507, July 1979.

[11] 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, November 1984.