Theoretical Analysis and Performance Limits of Noncoherent Sequence Detection of Coded PSK

Giulio Colavolpe, *Student Member, IEEE,* and Riccardo Raheli, *Member, IEEE*

***Abstract*—In this paper, a theoretical performance analysis of noncoherent sequence detection schemes recently proposed by the authors for combined detection and decoding of coded M-ary phase-shift keying (M-PSK) is presented. A method for the numerical evaluation of the pairwise error probability—for which no closed-form expressions exist—is described, the classical union bound is computed, and results are compared with computer simulations. An upper bound on this pairwise error probability is also presented. This upper bound may be effectively used for the definition of an equivalent distance, which may be useful in exhaustive searches for optimal codes. Using this bound, it is proven that, in the general coded case, the considered noncoherent decoding schemes perform as close as desired to an optimal coherent receiver when a phase memory parameter is sufficiently large. In the case of differentially encoded M-PSK, a simple expression of the asymptotic bit-error probability is derived, which is in agreement with simulations for high as well as low signal-to-noise ratio (SNR).**

***Index Terms*—Codecs, maximum-likelihood decoding, noncoherent sequence detection, phase-shift keying, Viterbi decoding.**

I. Introduction

In the technical literature, a growing interest has been recently shown toward improved noncoherent detection or decoding schemes [1]–[11]. As a general result of most of the above bibliographical references, the performance of noncoherent schemes based on extended windows of observation of the received signal improves, for increasing observation length and receiver complexity, and rapidly approaches that of optimal coherent receivers. Although this result has been mainly based on computer simulations, it has been analytically proven in the case of differentially encoded M-PSK [6], and, more recently, in the case of coded binary phase-shift keying (BPSK) [10], [11].

In [12], new noncoherent sequence detection schemes for combined detection and decoding of coded linear modulations have been proposed, which compare favorably with previously proposed solutions. This approach "moves beyond the crude block detection found in this area, to more natural, logical and optimal trellis-based detection schemes."[1] The starting point in [12] is the optimal noncoherent receiver which operates in the presence of a random phase rotation of the received signal, modeled as constant during the entire transmission, and additive white Gaussian noise (AWGN). Since optimal sequence detection requires a search of a path in a tree diagram, the required complexity increases exponentially with the duration of the transmission. Proper approximations are proposed in [12] in order to reduce the problem to a search of a path in a trellis diagram and realize simple suboptimal schemes based on the Viterbi algorithm. Besides being realizable, these schemes have the convenient feature of allowing us to remove the constant phase assumption and encompassing time-varying phase models. For increasing complexity, the performance of these receivers is shown to approach that of optimal coherent receivers by means of computer simulation [12]. A performance gain may be obtained with respect to existing schemes, such as those in [6] and [9], with acceptable levels of complexity—the tradeoff between complexity and performance being simply controlled by a parameter referred to as *implicit phase memory* and the number of trellis states [12]. As an example, assuming a value of implicit phase memory equal to the block length in [6], noncoherent sequence detection outperforms multiple-symbol differential detection [6], as shown in [12] by means of computer simulations. Moreover, noncoherent sequence detection has better performance with respect to the algorithm proposed in [9], whenever the decoding complexity is constrained to an affordable number of trellis states [12].

In this paper, we theoretically analyze the performance of a noncoherent sequence detection scheme proposed in [12] for coded M-PSK. We first resort to a numerical evaluation of the pairwise error probability, which admits a closed-form expression in very special cases only. In fact, in this case, a suitable decision variable may be expressed as an indefinite Hermitian quadratic form in a Gaussian nonzero mean vector, sometimes referred to as *noncentral*. Therefore, the method based on the residue theorem used in [13] for *central* quadratic forms cannot be used. Similar numerical methods are proposed in [14]–[16]. Based on the numerical evaluation of the pairwise error probability, we easily compute the classical union upper bound on the bit-error probability. This bound is found to be in excellent agreement with the simulation results presented in [12].

In [9], it was noted that optimal codes under coherent decoding are not necessarily optimal under noncoherent decoding. This phenomenon also occurs in the considered noncoherent schemes, as shown in Section IV. In principle, using the derived union bound, we could perform an exhaustive search for optimal codes. However, the required computational effort would be heavy because possible codes should be compared in terms of the entire curve of bit-error probability versus signal-to-noise

Manuscript received July 20, 1998; revised February 17, 2000. This work was supported by Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MURST), Italy. The material in this paper was presented in part at the IEEE International Conference on Communications (ICC'98), Atlanta, GA.

The authors are with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, 43100 Parma, Italy.

Communicated by E. Soljanin, Associate Editor for Coding Techniques.
Publisher Item Identifier S 0018-9448(00)04648-4.

[1]Quoted from an anonymous review of [12].

ratio (SNR), instead of a single parameter, such as the Euclidean distance which is sufficient in the case of coherent decoding. For this reason, we introduce an upper bound on the pairwise error probability that allows us to define a parameter, referred to as *equivalent distance*, which plays a role similar to that of the Euclidean distance in the case of coherent decoding. The minimum equivalent distance provides a comprehensive description of the performance of a specific code on a random-phase AWGN channel. The concept of equivalent distance has been previously proposed in the case of full response M-ary continuous phase modulations in [17], where the error probability, expressed in terms of the time-continuous waveform by means of the Marcum Q function, is approximated as a Gaussian Q function.

Using this equivalent distance we prove that, for any given code, the considered noncoherent receivers have a performance which approaches that of the optimal coherent receiver when the implicit phase memory parameter increases (in the following, we denote with N the implicit phase memory). Although when $N \to \infty$ a code which is optimal under coherent decoding is also optimal under noncoherent decoding, this is not necessarily true for finite values of N. In this case, a search for optimal codes may be performed using the proposed minimum equivalent distance.

For noncoherent sequence detection of differentially encoded M-PSK, we derive a simple expression of the asymptotic high-SNR bit-error probability which is in very good agreement with simulations for low SNR as well. Using this expression, we explicitly confirm the property of these noncoherent receivers of having a performance which approaches that of the optimal coherent receiver when the implicit phase memory parameter increases.

II. Signal Model and Noncoherent Sequence Detection

Assuming absence of intersymbol interference, the samples x_k at the output of a filter matched to the transmitted pulse may be expressed as [12]

$$x_k = c_k e^{j\theta} + n_k \tag{1}$$

in which, the code symbols $\{c_k\}$ belong to an M-PSK alphabet

$$\mathcal{A} \triangleq \{e^{j\phi_k} : \phi_k = \frac{2\pi m}{M}; m = 0, 1, \cdots, M-1\}$$

and are assumed to be derived from an information sequence $\{a_k\}$, composed of independent and identically distributed symbols belonging to an M'-ary alphabet, by means of some coding rule, θ is a constant phase shift introduced by the channel, and n_k are independent, identically distributed, zero-mean, complex, Gaussian random variables with independent real and imaginary components, each with variance $\sigma^2 = (N_0/2E_S)$, where N_0 is the one-sided power spectral density of the bandpass noise and E_S is symbol energy.

In [12], two coding rules for which symbols a_k and c_k belong to the same M-ary alphabet $\mathcal{A}$ $(M' = M)$ are explicitly considered in the simulations: differential encoding, specified by the recursion $c_k = c_{k-1}a_k$, and M-ary convolutional encoding, with encoder structure described in [9]. In the latter case, encoder structures which consist of only one shift register with code rate $1/\eta$, where η is the number of code symbols per information symbol, are considered.[2] The number of code states is $S_c = M^{\nu-1}$, in which ν denotes the code constraint length.

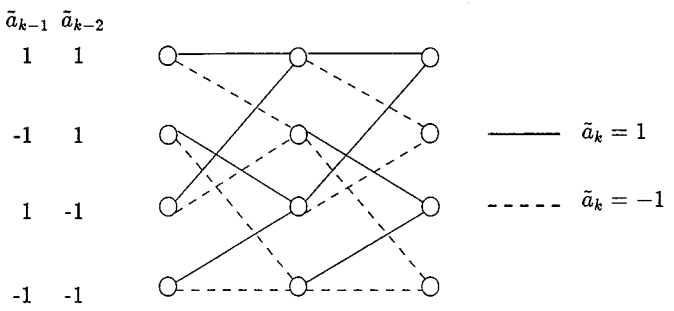

Fig. 1. Decoder trellis for differentially encoded BPSK and $N = 4$.

In the numerical examples, only the described codes are considered. However, for generality, this paper addresses the case $M' \neq M$—the case $M' < M$ being of interest in some applications such as trellis-coded modulations (TCM).

The branch metrics of a Viterbi algorithm which realizes the considered noncoherent sequence detection or decoding scheme are [12]

$$\lambda_k = \mathrm{Re}\left\{\sum_{i=1}^{N-1} x_k x_{k-i}^* \tilde{c}_k^* \tilde{c}_{k-i}\right\} \tag{2}$$

in which, the parameter N plays the role of an implicit phase memory and $\{\tilde{c}_k\}$ is a hypothetical code sequence. Given the coding rule, these branch metrics may be expressed in terms of the hypothetical information sequence $\{\tilde{a}_k\}$, uniquely associated to the code sequence $\{\tilde{c}_k\}$. As an example, in the case of differential encoding, the branch metrics (2) become

$$\lambda_k = \mathrm{Re}\left\{\sum_{i=1}^{N-1} x_k x_{k-i}^* \prod_{m=0}^{i-1} \tilde{a}_{k-m}^*\right\} \tag{3}$$

Even using small values of N (a few units) a performance very close to that of coherent detection or decoding may be obtained [12].

According to (2), a trellis state may be properly defined in terms of γ information symbols as

$$\delta_k \triangleq (\tilde{a}_{k-1}, \tilde{a}_{k-2}, \cdots, \tilde{a}_{k-\gamma})$$

in which $\gamma = N-2$ for differential encoding and $\gamma = N/\eta + \nu - 2$ for convolutional encoding, and the resulting number of states is $S = M^\gamma$ [12]. In fact, in the case of differential encoding, it may be observed that the branch metrics (3) depend on the current symbol $\tilde{a}_k$ and the previous $N-2$ information symbols. The trellis structure in the case of differentially encoded BPSK $(M = 2)$ and $N = 4$ is shown in Fig. 1 ($S = 2^{N-2} = 4$ in this case). We remark that the noncoherent nature of the considered schemes is due to the presence of terms of the form $x_i x_j^*$ in the branch metrics. As a consequence, a constant phase shift θ does not affect the receiver performance—in the following derivations, $\theta = 0$ is assumed.

[2]Other rates are possible using multiple shift registers.

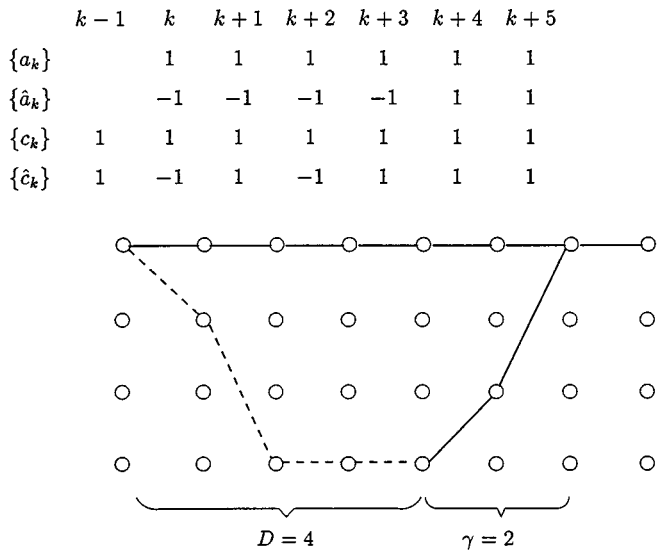

Fig. 2. An error event for differentially encoded BPSK and $N = 4$.

III. Numerical Evaluation of the Union Bound

The classical union bound[3] on the probability of bit error P_b has expression [18]

$$P_b \leq \frac{1}{\log_2 M} \sum_{\boldsymbol{a}} P(\boldsymbol{a}) \sum_{\hat{\boldsymbol{a}} \neq \boldsymbol{a}} b(\boldsymbol{a}, \hat{\boldsymbol{a}}) P(\boldsymbol{a} \to \hat{\boldsymbol{a}}) \tag{4}$$

in which $\boldsymbol{a} \triangleq \{a_i\}$ and $\hat{\boldsymbol{a}} \triangleq \{\hat{a}_i\}$ denote the information sequences corresponding to the correct and erroneous paths, respectively, $b(\boldsymbol{a}, \hat{\boldsymbol{a}})$ is the number of bit errors entailed by the error event $(\boldsymbol{a}, \hat{\boldsymbol{a}})$, $P(\boldsymbol{a} \to \hat{\boldsymbol{a}})$ is the pairwise error probability, and $P(\boldsymbol{a})$ is the *a priori* probability of sequence $\boldsymbol{a}$. Assuming a duration of the error event of $D + \gamma$ symbol intervals, the correct and erroneous paths differ at most by D symbols. As an example, in the case of a differentially encoded BPSK with $N = 4$ whose trellis diagram is shown in Fig. 1, an error event of duration six symbol intervals ($D = 4$ and $\gamma = N - 2 = 2$) is shown in Fig. 2, along with the information and code sequences associated with the correct and erroneous paths.

If proper symmetry conditions on the considered code, receiver metrics, and bit labeling are verified, the inner sum in (4) may be independent of the transmitted sequence $\boldsymbol{a}$. In the case of Gray labeling, the differential and convolutional codes considered in the examples satisfy these conditions under the general class of noncoherent decoders which maximize an arbitrary likelihood function expressed in terms of the output of a correlator front-end [19]. The considered noncoherent schemes belong to this class. This implies that, in the performance evaluation, we may assume that a specific sequence $\boldsymbol{a}$ is transmitted.

[3]The terminology "union bound" is used here, though more appropriate for the error-even probability.

In this case, the classical union bound (4) on the probability of bit error P_b simplifies to

$$P_b \leq \frac{1}{\log_2 M} \sum_{\hat{\boldsymbol{a}} \neq \boldsymbol{a}} b(\boldsymbol{a}, \hat{\boldsymbol{a}}) P(\boldsymbol{a} \to \hat{\boldsymbol{a}}) \tag{5}$$

in which the transmitted information sequence $\boldsymbol{a}$ is arbitrarily selected. For example, for both the described differential and convolutional encoding rules, we may assume that the transmitted information sequence is given by $a_i = 1, \forall i$, and the corresponding code sequence is such that $c_i = 1, \forall i$. The theoretical results presented in this paper are relative to the general case (4); however, examples and numerical results are given for codes satisfying these symmetry conditions.

The pairwise error probability $P(\boldsymbol{a} \to \hat{\boldsymbol{a}})$ is the probability that the sum of the branch metrics $\hat{\lambda}_i$ relative to the erroneous path exceeds the sum of the metrics λ_i on the correct path. An error event beginning at time k and involving D information symbols ends at time $k + D + \gamma$. Hence, $P(\boldsymbol{a} \to \hat{\boldsymbol{a}})$ may be expressed in terms of the decision variable

$$y \triangleq \sum_{i=k}^{k+D+\gamma-1} \hat{\lambda}_i - \lambda_i.$$

In the steady state, the statistics of this decision variable are independent of k. Without loss of generality, we may select a suitable value of k in order to simplify the notation. Specifically, we use a value such that the decision variable y only depends on signal samples x_i with index $i \geq 0$ (for example, $k = N - 1$ for differential encoding and $k = N/\eta - 1$ for convolutional encoding). With this assumption, y may be expressed in terms of the Gaussian vector $\boldsymbol{x} \triangleq (x_0, x_1, \cdots, x_{P-1})^T$ of P successive matched filter outputs. The actual value of P depends on code characteristics, such as rate and constraint length. For example, see (6) at the bottom of this page.

Conditionally on a specific transmitted sequence, the Gaussian vector $\boldsymbol{x}$ has independent complex Gaussian components $x_i = c_i + n_i$ with mean c_i and variance $2\sigma^2$. Its conditional mean vector and covariance matrix are

$$\boldsymbol{\eta}_{\boldsymbol{x}} \triangleq E\{\boldsymbol{x}\} = (c_0, c_1, \cdots, c_{P-1})^T \tag{7}$$

$$\boldsymbol{C}_{\boldsymbol{x}} \triangleq E\{(\boldsymbol{x} - \boldsymbol{\eta}_{\boldsymbol{x}})(\boldsymbol{x} - \boldsymbol{\eta}_{\boldsymbol{x}})^H\} = 2\sigma^2 \boldsymbol{I} \tag{8}$$

where $\boldsymbol{I}$ denotes the identity matrix and $[\cdot]^H$ is the transpose conjugate operator.

For a given error event $(\boldsymbol{a}, \hat{\boldsymbol{a}})$, the random variable y may be expressed as a quadratic form and the pairwise error probability is given by

$$P(\boldsymbol{a} \to \hat{\boldsymbol{a}}) = P\{y > 0\} = P\{\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x} > 0\} \tag{9}$$

where the Hermitian $P \times P$ matrix $\boldsymbol{A}$ has entries

$$A_{ij} = \hat{c}_i \hat{c}_j^* - c_i c_j^* \tag{10}$$

$$P = \begin{cases} 2N + D - 3, & \text{for differentially encoded } M\text{-PSK} \\ \eta(D + 2N/\eta + \nu - 3), & \text{for convolutionally encoded } M\text{-PSK.} \end{cases} \tag{6}$$

in which $\hat{c}_i$ is the ith code symbol associated with the considered erroneous sequence. Despite the fact that $\boldsymbol{x}$ and $\boldsymbol{A}$ individually depend on the transmitted sequence, for the considered codes the quadratic form (9) does not, in agreement with the mentioned symmetry conditions.

In order to describe the structure of matrix $\boldsymbol{A}$, it is necessary to introduce another parameter, denoted by D', which represents the maximum number of erroneous code symbols entailed by an error event of length $D+\gamma$. With this definition, matrix $\boldsymbol{A}$, being Hermitian, is uniquely defined by D' columns only. As an example, in the special case of an error event such as that shown in Fig. 2, for which $D=4$, $D'=3$, and for $N=4$ (hence, $k=3$ and $P=9$), noting that $\hat{c}_i = c_i$ for $i<k$ and $i>k+D'-1$, the matrix $\boldsymbol{A}$ may be expressed as

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{B}^H & \boldsymbol{C} & \boldsymbol{D}^H \\ \boldsymbol{O} & \boldsymbol{D} & \boldsymbol{O} \end{pmatrix} \tag{11}$$

where $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ are defined at the bottom of this page and $\boldsymbol{O}$ is a 3×3 zero-entry matrix.

This matrix may be diagonalized as $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{M}\boldsymbol{Q}^{-1}$, where $\boldsymbol{M}\triangleq \operatorname{diag}(\mu_i)$ is the diagonal eigenvalue matrix of $\boldsymbol{A}$ (μ_i denotes the ith eigenvalue), $\boldsymbol{Q}$ is unitary (i.e., $\boldsymbol{Q}^{-1}=\boldsymbol{Q}^H$) and its columns are the eigenvectors of $\boldsymbol{A}$. Since $\boldsymbol{A}$ is Hermitian, its eigenvalues are real. As shown in Appendix A, if μ is a nonzero eigenvalue of $\boldsymbol{A}$, $-\mu$ is also an eigenvalue; hence, the number of nonzero eigenvalues is even. The quadratic form may be expressed as

$$y=\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^H\boldsymbol{Q}\boldsymbol{M}\boldsymbol{Q}^{-1}\boldsymbol{x}=\boldsymbol{z}^H\boldsymbol{M}\boldsymbol{z}=\sum_{i=1}^{2E}\mu_i|z_i|^2 \tag{12}$$

in which $\boldsymbol{z}\triangleq\boldsymbol{Q}^{-1}\boldsymbol{x}=\boldsymbol{Q}^H\boldsymbol{x}$ and $2E$ denotes the even number of nonzero eigenvalues.

The random vector $\boldsymbol{z}$ is Gaussian with mean and covariance given by

$$\begin{aligned}\boldsymbol{\eta_z}&\triangleq E\{\boldsymbol{z}\}=\boldsymbol{Q}^{-1}\boldsymbol{\eta_x}\\ &=\boldsymbol{Q}^H(c_0,c_1,\cdots,c_{P-1})^T=(\eta_{z_1},\eta_{z_2},\cdots,\eta_{z_P})^T\\ \boldsymbol{C_z}&\triangleq E\{(\boldsymbol{z}-\boldsymbol{\eta_z})(\boldsymbol{z}-\boldsymbol{\eta_z})^H\}\\ &=\boldsymbol{Q}^H E\{(\boldsymbol{x}-\boldsymbol{\eta_x})(\boldsymbol{x}-\boldsymbol{\eta_x})^H\}\boldsymbol{Q}=2o^2\boldsymbol{I}.\end{aligned} \tag{13}$$

Therefore, z_i are complex, Gaussian, independent random variables with nonzero mean and $|z_i|^2$ have a noncentral chi-square distribution with two degrees of freedom [20], [21]. Letting $y_i\triangleq\mu_i|z_i|^2$, its mean, variance, and bilateral Laplace transform of the probability density function are, respectively, for $\mu_i\neq 0$ [20], [21]

$$\eta_{y_i}=E\{y_i\}=\mu_i(2\sigma^2+|\eta_{z_i}|^2) \tag{14}$$

$$\sigma^2_{y_i}=\mu_i^2(4\sigma^4+4\sigma^2|\eta_{z_i}|^2) \tag{15}$$

$$\Psi_{y_i}(s)=-\frac{p_i}{s-p_i}\exp\left\{\frac{ss_ip_i}{s-p_i}\right\} \tag{16}$$

where $s_i\triangleq\mu_i|\eta_{z_i}|^2$ and $p_i\triangleq -1/2\mu_i\sigma^2$. Due to the independence of the random variables y_i, the bilateral Laplace transform of the probability density function of y may be expressed as

$$\Psi_y(s)=\prod_{i=1}^{2E}\frac{-p_i}{s-p_i}\exp\left\{\frac{ss_ip_i}{s-p_i}\right\}. \tag{17}$$

Since a closed-form expression of the probability density function of y does not exist, for the evaluation of $P(\boldsymbol{a}\to\hat{\boldsymbol{a}})$ one could use $\Psi_y(s)$ and the residue theorem as in [13]

$$\begin{aligned}P(\boldsymbol{a}\to\hat{\boldsymbol{a}})&=1-P\{y\le 0\}=1-\frac{1}{j2\pi}\int_{\epsilon-j\infty}^{\epsilon+j\infty}\frac{\Psi_y(s)}{s}\,ds\\ &=1+\sum\operatorname{Res}\left[\frac{\Psi_y(s)}{s}\right]\Bigg|_{\mathrm{RHP}}\end{aligned} \tag{18}$$

where $\epsilon>0$ is a proper small constant and the residues of the singularities in the right half-plane (RHP) are considered, with the exception of the pole for $s=0$. The property of the eigenvalues of matrix $\boldsymbol{A}$, shown in Appendix A, of being in equal and opposite real pairs allows us to conclude that singularities p_i of the function $\Psi_y(s)$ share the same property. Unfortunately, these singularities are essential and their residues may not be expressed in a closed form. Several numerical methods are known in the literature for the computation of these residues (see [14]–[16], and references therein). Among the various options, we calculated directly the sum of the residues by numerical integration on a sufficiently large square path including all singularities in the right half-plane.

In the case of the considered codes for which the described symmetry conditions for the validity of (5) hold, it is possible to truncate the union bound considering only error events with $D\le D_0$, possibly selecting D_0 in such a way that the contribution of error events with $D>D_0$ is negligible. In the case of differentially encoded QPSK, Fig. 3 shows the truncated upper

$$\boldsymbol{B}\triangleq\begin{pmatrix} c_0(\hat{c}_k^*-c_k^*) & c_0(\hat{c}_{k+1}^*-c_{k+1}^*) & c_0(\hat{c}_{k+2}^*-c_{k+2}^*)\\ c_1(\hat{c}_k^*-c_k^*) & c_1(\hat{c}_{k+1}^*-c_{k+1}^*) & c_1(\hat{c}_{k+2}^*-c_{k+2}^*)\\ c_2(\hat{c}_k^*-c_k^*) & c_2(\hat{c}_{k+1}^*-c_{k+1}^*) & c_2(\hat{c}_{k+2}^*-c_{k+2}^*)\end{pmatrix}$$

$$\boldsymbol{C}\triangleq\begin{pmatrix} 0 & \hat{c}_{k+1}^*\hat{c}_k-c_{k+1}^*c_k & \hat{c}_{k+2}^*\hat{c}_k-c_{k+2}^*c_k\\ \hat{c}_k^*\hat{c}_{k+1}-c_k^*c_{k+1} & 0 & \hat{c}_{k+2}^*\hat{c}_{k+1}-c_{k+2}^*c_{k+1}\\ \hat{c}_k^*\hat{c}_{k+2}-c_k^*c_{k+2} & \hat{c}_{k+1}^*\hat{c}_{k+2}-c_{k+1}^*c_{k+2} & 0\end{pmatrix}$$

$$\boldsymbol{D}\triangleq\begin{pmatrix} c_{P-3}(\hat{c}_k^*-c_k^*) & c_{P-3}(\hat{c}_{k+1}^*-c_{k+1}^*) & c_{P-3}(\hat{c}_{k+2}^*-c_{k+2}^*)\\ c_{P-2}(\hat{c}_k^*-c_k^*) & c_{P-2}(\hat{c}_{k+1}^*-c_{k+1}^*) & c_{P-2}(\hat{c}_{k+2}^*-c_{k+2}^*)\\ c_{P-1}(\hat{c}_k^*-c_k^*) & c_{P-1}(\hat{c}_{k+1}^*-c_{k+1}^*) & c_{P-1}(\hat{c}_{k+2}^*-c_{k+2}^*)\end{pmatrix}$$

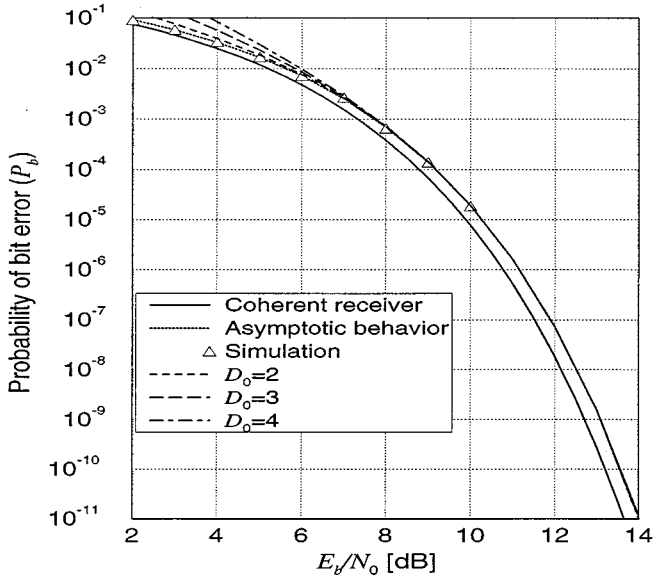

Fig. 3. Upper bound for differentially encoded QPSK (receiver with $N = 4$), for various values of D_0 and comparison with simulation results. The asymptotic high-SNR performance derived in Section V is also shown.

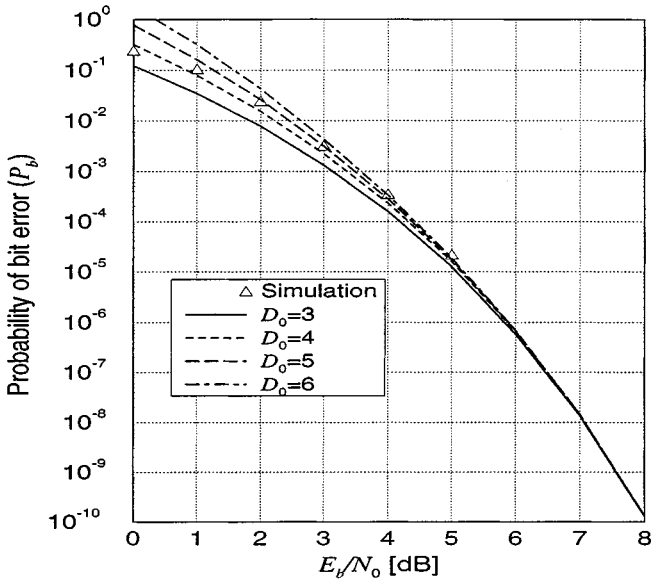

Fig. 4. Upper bound for QPSK with the convolutional code of rate $1/2$ and constraint length $\nu = 3$ considered in [12], for various values of D_0, and comparison with simulation results (receiver with $N = 8$).

bound for the receiver with $N = 4$ and compares it with the results of computer simulations. We may observe that the analysis is in very good agreement with simulation. In the figure, various values of parameter D_0 are considered. The asymptotic behavior for high SNR is not affected by the value of D_0, confirming a good convergence of the union bound. In the figure, the asymptotic high-SNR performance based on the most probable error events, analytically derived in closed form in Section V, is also shown. Fig. 4 shows the truncated upper bound for QPSK with the convolutional code of rate $1/2$, constraint length $\nu = 3$, and generators[4] 133, 231 considered in [12] and compares it with computer simulations. This figure confirms the validity of the conclusions previously derived in the absence of coding and shows that the analysis is again in good agreement with simulation.

IV. Upper Bound on the Pairwise Error Probability

The described method for the computation of the truncated union bound cannot be efficiently used in the search for optimal codes for noncoherent decoding because, for each possible code, the entire curve of bit-error probability versus SNR has to be derived. In this section, we introduce an upper bound on the pairwise error probability, in order to define a parameter which could play a role similar to that of the Euclidean distance in the case of coherent detection.

The Chernoff bound on the pairwise error probability is

$$P(\boldsymbol{a} \to \hat{\boldsymbol{a}}) = P\{y \geq 0\} \leq E\{e^{\zeta y}\} \tag{19}$$

for each value of $\zeta \geq 0$ [20]. The tightest upper bound may be obtained selecting $\zeta = \zeta_0$, where ζ_0 minimizes the right side of (19).

[4]Base-4 representation.

By properly ordering the eigenvalues of $\boldsymbol{A}$, we may express (12) as

$$y = \sum_{i=1}^{2E} \mu_i |z_i|^2 = \sum_{i=1}^{E} \mu_{2i-1}(|z_{2i-1}|^2 - |z_{2i}|^2) = \sum_{i=1}^{E} w_i \tag{20}$$

in which $w_i \triangleq \mu_{2i-1}(|z_{2i-1}|^2 - |z_{2i}|^2)$. Using (14) and (15), we have

$$\eta_{w_i} = E\{w_i\} = \mu_{2i-1}\left(|\eta_{z_{2i-1}}|^2 - |\eta_{z_{2i}}|^2\right) \tag{21}$$

$$\sigma_{w_i}^2 = 8\sigma^4\mu_{2i-1}^2 + 4\sigma^2\mu_{2i-1}^2\left(|\eta_{z_{2i-1}}|^2 + |\eta_{z_{2i}}|^2\right). \tag{22}$$

The moment generating function $E\{e^{\zeta y}\}$ [20], after some straightforward manipulations, may be written as

$$E\{e^{\zeta y}\} = \int_{-\infty}^{\infty} e^{\zeta y} p_y(y)\, dy = \Psi_y(-\zeta) = \prod_{i=1}^{E} \frac{1}{1 - 4\mu_{2i-1}^2\sigma^4\zeta^2} \exp\left\{\frac{\eta_{w_i}\zeta + \frac{s_{w_i}^2}{2}\zeta^2}{1 - 4\mu_{2i-1}^2\sigma^4\zeta^2}\right\} \tag{23}$$

in which

$$s_{w_i}^2 \triangleq \sigma_{w_i}^2 - 8\mu_{2i-1}^2\sigma^4 = 4\sigma^2\mu_{2i-1}^2\left(|\eta_{z_{2i-1}}|^2 + |\eta_{z_{2i}}|^2\right). \tag{24}$$

Letting $\mu_M \triangleq \max_i\{|\mu_i|\}$, the integral which defines the moment generating function converges in the interval

$$\zeta \in \left[-\frac{1}{2\mu_M\sigma^2}, \frac{1}{2\mu_M\sigma^2}\right]$$

only. In order to minimize $E\{e^{\zeta y}\}$, we note that for values of ζ such that $4\mu_M^2\sigma^4\zeta^2 \ll 1$, i.e., $\zeta \ll (1/2\mu_M\sigma^2)$, the following approximation holds:

$$E\{e^{\zeta y}\} \simeq \prod_{i=1}^{E} \exp\left\{\eta_{w_i}\zeta + \frac{s_{w_i}^2}{2}\zeta^2\right\} = \exp\left\{\eta_y\zeta + \frac{s_y^2}{2}\zeta^2\right\} \tag{25}$$

where $\eta_y = \sum_{i=1}^{E}\eta_{w_i}$ and $s_y^2 \triangleq \sum_{i=1}^{E} s_{w_i}^2$. The minimum of approximation (25) is achieved for $\zeta_0 = -(\eta_y/s_y^2)$ and equals $\exp\{-(\eta_y^2/2s_y^2)\}$. This result may be considered an approximate minimum of $E\{e^{\zeta y}\}$ if the inequality $\zeta_0 \ll (1/2\mu_M\sigma^2)$ is verified, i.e., recalling (21), (22), and the definition of $s_{w_i}^2$, if

$$2\sigma^2\zeta_0 = -\frac{2\sigma^2\eta_y}{s_y^2} = -\frac{\sum_{i=1}^{E}\mu_{2i-1}\left(\left|\eta_{z_{2i-1}}\right|^2 - \left|\eta_{z_{2i}}\right|^2\right)}{2\sum_{i=1}^{E}\mu_{2i-1}^2\left(\left|\eta_{z_{2i-1}}\right|^2 + \left|\eta_{z_{2i}}\right|^2\right)} \ll \frac{1}{\mu_M}. \tag{26}$$

This condition is independent of the SNR and depends on matrix $\boldsymbol{A}$ only. We verified that (26) is satisfied for every error event we considered. The validity of this condition will be explicitly addressed in a following paragraph.

Based on the above result, the approximate Chernoff upper bound on the pairwise error probability is

$$P(\boldsymbol{a}\to\hat{\boldsymbol{a}}) = P\{y \ge 0\} \le E\left\{\exp\left[-\frac{\eta_y}{s_y^2}y\right]\right\} \simeq \exp\left\{-\frac{\eta_y^2}{2s_y^2}\right\} = \exp\left\{-\frac{d_{\rm eq}^2}{8\sigma^2}\right\} \tag{27}$$

in which

$$d_{\rm eq} \triangleq \frac{2|\eta_y|\sigma}{s_y} = \frac{\left|\sum_{i=1}^{E}\mu_{2i-1}\left(\left|\eta_{z_{2i-1}}\right|^2 - \left|\eta_{z_{2i}}\right|^2\right)\right|}{\sqrt{\sum_{i=1}^{E}\mu_{2i-1}^2\left(\left|\eta_{z_{2i-1}}\right|^2 + \left|\eta_{z_{2i}}\right|^2\right)}}. \tag{28}$$

This parameter can be calculated without explicitly diagonalizing matrix $\boldsymbol{A}$. In fact, expressing $\boldsymbol{x}$ as a sum of $\boldsymbol{\eta_x}$ and noise terms, it may be shown that ([22, Appendix A])

$$d_{\rm eq} = \frac{\left|\boldsymbol{\eta_x}^H\boldsymbol{A}\boldsymbol{\eta_x}\right|}{\sqrt{\boldsymbol{\eta_x}^H\boldsymbol{A}^2\boldsymbol{\eta_x}}}. \tag{29}$$

Parameter $d_{\rm eq}$ determines the asymptotic high-SNR behavior of the pairwise error probability and plays a role similar to that of the Euclidean distance in coherent decoding; for this reason, it is referred to as *equivalent distance* of the error event. We remark that this approach may be applied whenever the decision variable is a quadratic form in a nonzero mean Gaussian vector. As an example, an equivalent distance may be similarly defined in the case of the receiver proposed in [9].

A tighter upper bound may be obtained in a more involved manner (see [22, Appendix C] for details), under the same assumption (26) (i.e., $4\mu_M^2\sigma^4\zeta_0^2 \ll 1$). This bound is identical to the modified Chernoff bound [23], although the derivation is here different. The bound reads

$$P(\boldsymbol{a}\to\hat{\boldsymbol{a}}) = P\{y \ge 0\} \le \frac{1}{\sqrt{2\pi}s_y|\zeta_0|}\exp\left\{-\frac{d_{\rm eq}^2}{8\sigma^2}\right\} = \frac{1}{\sqrt{2\pi}\frac{d_{\rm eq}}{2\sigma}}\exp\left\{-\frac{d_{\rm eq}^2}{8\sigma^2}\right\} \simeq Q\left(\frac{d_{\rm eq}}{2\sigma}\right) \tag{30}$$

where

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}\,dt$$

is the Gaussian Q function and we used the property that, for $x > 3$

$$Q(x) \simeq \frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}} \le e^{-\frac{x^2}{2}}. \tag{31}$$

The bound (30) is evidently tighter than (27) because they differ for a multiplying factor only, which is less than 1 for $\frac{d_{\rm eq}}{2\sigma} > \frac{1}{\sqrt{2\pi}}$. This condition is always met for pairwise error probabilities less than 0.35. In addition, if $\frac{d_{\rm eq}}{2\sigma} > 3$, the bound may be approximated by the Gaussian Q function as suggested by (31). This condition is verified for pairwise error probabilities less then 10^{-3}. In the following, we denote this approximate bound as *Gaussian bound*.

We now prove an important property of the considered noncoherent receivers. We begin with a property of the equivalent distance and one of the proposed bounds.

Lemma 1: The equivalent distance of any error event $(\boldsymbol{a},\hat{\boldsymbol{a}})$ satisfies

$$\lim_{N\to\infty} d_{\rm eq} = d_E \tag{32}$$

in which d_E is the Euclidean distance of the error event in coherent decoding.

Proof: In Appendix B, it is shown that the equivalent distance of any error event may be expressed as

$$d_{\rm eq}^2 = \left(1 - \frac{D'}{P}\right)d_E^2 + \frac{1}{P}d_x^2 \tag{33}$$

in which

$$d_x^2 \triangleq 2\sum_{i=k}^{k+D'-1}\sum_{j=k}^{i-1}\mathrm{Re}[(1 - c_i^*c_j\hat{c}_i\hat{c}_j^*)]. \tag{34}$$

When $N\to\infty$, also $P\to\infty$ and $d_{\rm eq}^2 \to d_E^2$. □

Lemma 2: When $N \to \infty$, the approximate Chernoff and Gaussian upper bounds are asymptotically exact. The Gaussian bound is also asymptotically tight.

Proof: See Appendix B for a proof that both approximate upper bounds are asymptotically exact. Since the Gaussian bound asymptotically coincides with the pairwise error probability in coherent decoding (Lemma 1), it is also asymptotically tight. □

Since these lemmas hold for every error event, the following theorems are easily proven.

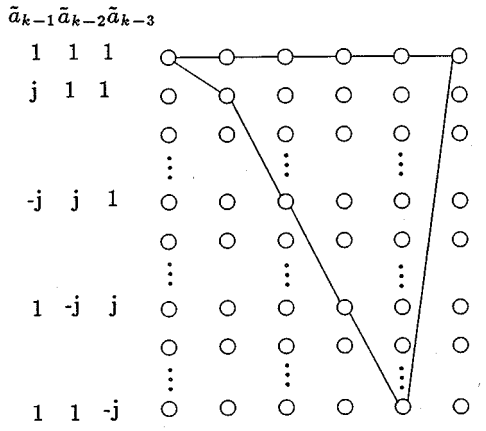

Fig. 5. Most probable error event $\mathcal{E}_0$ for differentially encoded QPSK with $N = 5$.

Theorem 1: The union bound for coherent decoding and that for the considered noncoherent decoding schemes coincide when $N \to \infty$.

Theorem 2: When $N \to \infty$, a code which is optimal under coherent decoding is also optimal under noncoherent decoding. For finite values of N, this is not necessarily true.

Based on Theorems 1 and 2, we may conclude that any coded scheme may be decoded noncoherently with performance as close as desired to that of an optimal coherent receiver. However, for finite complexity a search for optimal codes may be useful. Since for high SNR, the performance is determined by the error events with minimum equivalent distance, this parameter may be used in the search for optimal codes.

In order to give examples, we now consider differentially encoded M-PSK. In this case the mentioned symmetry conditions hold; hence, we may assume the symbol sequence $c_i = 1, \forall i$, is transmitted. By direct computation of the equivalent distance or pairwise error probability, as described in Section III, it is easy to verify that for $N \geq 4$, the most probable error events starting at time k are $\mathcal{E}_0$ and $\mathcal{E}_0'$, characterized by

$$\hat{a}_k = a_k e^{j(2\pi/M)}, \hat{a}_{k+1} = a_k e^{-j(2\pi/M)}$$

and

$$\hat{a}_k = a_k e^{-j(2\pi/M)}, \hat{a}_{k+1} = a_k e^{j(2\pi/M)}$$

respectively ($D = 2$ and $P = 2N - 1$, according to (6)). In the special case of $M = 2$ (BPSK), the error events $\mathcal{E}_0$ and $\mathcal{E}_0'$ coincide. The error event $\mathcal{E}_0$ is shown in Fig. 5 for $M = 4$ (QPSK) and $N = 5$, along with the relevant trellis states.

The complete list of principal error events for differentially encoded QPSK with $N = 5$ is shown in Table I along with the relative equivalent distances. A minimum distance of 1.33333 is exhibited by error events $\mathcal{E}_0$ and $\mathcal{E}_0'$. There exist two second most probable error events with distance 1.78885. At the bottom of the table, we show a less frequent error event which is considered in the following figure.

In Fig. 6, the exact pairwise error probability, the Gaussian, and the Chernoff bound are shown for three typical error events listed in Table I. In this figure, the minimum-distance error event

TABLE I
LIST OF ERROR EVENTS WITH INCREASING EQUIVALENT DISTANCE, FOR DIFFERENTIALLY ENCODED QPSK WITH $N = 5$. THE CONSTELLATION POINTS $\{1, j, -1, -j\}$ ARE DENOTED BY $\{0, 1, 2, 3\}$, RESPECTIVELY

$\{\hat{a}_k\}$	Equivalent distance
1 3 3 1	1.33333
1 0 3 3 0 1	1.78885
1 3 1 3 3 1 3 1	1.80907
1 3 0 1 3 3 1 0 3 1	1.82574
1 3 0 0 1 3 3 1 0 0 3 1	1.83973
2 2	1.88562
1 2 1 3 2 3	1.89737
1 3 3 1 3 1 1 3	1.90693
1 3 0 3 1 3 1 0 1 3	1.91485
1 3 0 0 3 1 3 1 0 0 1 3	1.92154
⋮	⋮
3 3 2	2.23607
⋮	⋮
$0 \to 1, 1 \to j, 2 \to -1, 3 \to -j$	

$\mathcal{E}_0$ is considered, along with one of the two second most probable error events (denoted by $\mathcal{E}_1$), and the least probable error events listed in Table I (denoted by $\mathcal{E}_2$). This figure confirms that the Gaussian bound is more accurate than the derived Chernoff bound, as predicted.

A list similar to that in Table I is shown in Table II for convolutionally encoded QPSK with $N = 8$. We note that the minimum equivalent distance is increased, with respect to differentially encoded QPSK, due to the effect of the convolutional code.

As a final remark, we note that Theorems 1 and 2, derived in this section using the asymptotic property of the equivalent distance (Lemma 1) and tightness of the modified Chernoff bound (Lemma 2), may be also derived using an asymptotic property of the eigenvalues of matrix $\boldsymbol{A}$. This is shown in Appendix C, where the Marcum Q function and its asymptotic behavior described in Section V are also used. This appendix may be viewed as an alternative simplified proof of the theorems. We decided to

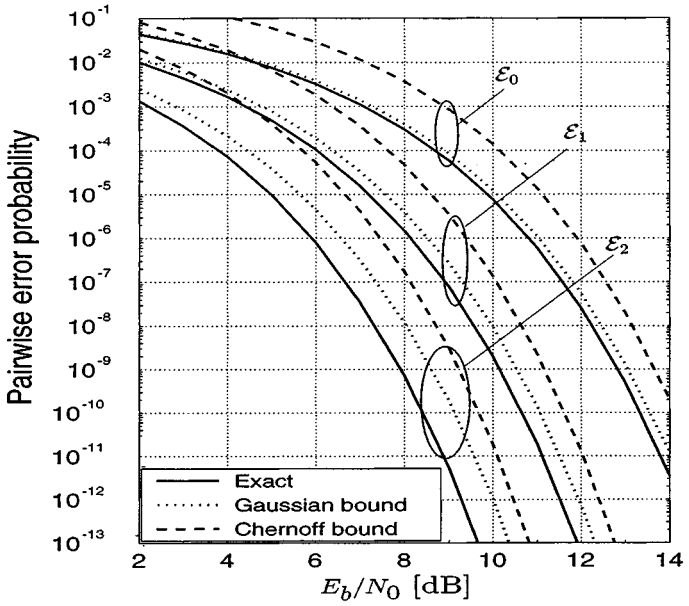

Fig. 6. Exact expression, Gaussian and Chernoff bound on the pairwise error probability of three typical error events for differentially encoded QPSK with $N = 5$. $\mathcal{E}_0$ is one of the two most probable error events ($d_{\text{cq}} = 1.33333$), $\mathcal{E}_1$ is one of the two second most probable error events ($d_{\text{cq}} = 1.78885$), and $\mathcal{E}_2$ is the least probable error event listed in Table I ($d_{\text{cq}} = 2.23607$).

emphasize the approach based on the modified Chernoff bound because it allows us to introduce the important tool referred to as equivalent distance.

V. Expression of P_b for Differentially Encoded M-PSK

In this section, we derive a closed-form expression of the asymptotic bit-error probability for noncoherent sequence detection of differentially encoded M-PSK. Since in this case the mentioned symmetry conditions hold, we may assume the symbols $c_i = 1, \forall i$, are transmitted.

As already mentioned in Section IV, the most probable error events starting at time k are $\mathcal{E}_0$ and $\mathcal{E}_0'$, with $\mathcal{E}_0$ characterized by $\hat{a}_k = a_k e^{j(2\pi/M)}$, $\hat{a}_{k+1} = a_k e^{-j(2\pi/M)}$ ($D = 2$ and $P = 2N-1$). Assuming Gray mapping, the number of corresponding errored bits is 2. The asymptotic bit-error probability depends on the pairwise error probability of these error events only. In the following, we consider only $\mathcal{E}_0$ because the pairwise error probabilities corresponding to $\mathcal{E}_0$ and $\mathcal{E}_0'$ are equal for symmetry.

In the case of error event $\mathcal{E}_0$, $D' = 1$ and matrix $\boldsymbol{A}$ has nonzero elements in the middle row and column only (i.e., those with index $N-1$). This row is $(a, \cdots, a, 0, a, \cdots, a)$, where $a \triangleq e^{j(2\pi/M)} - 1$, and the zero element is in the middle position (obviously, elements of the middle column are determined because $\boldsymbol{A}$ is Hermitian). In this case, there are two nonzero eigenvalues

$$\mu_{1,2} = \pm\sqrt{P-1}|a| \tag{35}$$

with the following corresponding eigenvectors:

$$\boldsymbol{q}_1 = (b\alpha, \cdots, b\alpha, \alpha, b\alpha, \cdots, b\alpha)^T \tag{36}$$

$$\boldsymbol{q}_2 = (-b\alpha, \cdots, -b\alpha, \alpha, -b\alpha, \cdots, -b\alpha)^T \tag{37}$$

TABLE II
List of Error Events with Increasing Equivalent Distance, for QPSK with the Convolutional Code of Rate $1/2$ and Constraint Length $\nu = 3$ Considered in [12] ($N = 8$). The Constellation Points $\{1, j, -1, -j\}$ Are Denoted by $\{0, 1, 2, 3\}$, Respectively

$\{\hat{a}_k\}$	Equivalent distance
1 3	3.51978
2 1 2 3	3.57071
1 1 3 3	3.70809
2 2	3.73497
2 1 3 2 3 1	3.75983
2	3.90868
2 1 3 3 2 2 0 1 2 2 0 3 2 3 1 1	3.93170
1 3 3 1	3.96862
1 0 1 3 0 3	3.97148
1 1 2 3 2 1 3 1 2 3 1 3 3 3 2 1	3.97387
⋮	⋮
$0 \to 1,\ 1 \to j,\ 2 \to -1,\ 3 \to -j$	

in which

$$b \triangleq \frac{a^*}{|a|}\frac{1}{\sqrt{P-1}}$$

and α is a normalizing factor such that $|\alpha|^2 = \frac{1}{2}$.

From (20), $y = \mu_1|z_1|^2 + \mu_2|z_2|^2 = \mu_1(|z_1|^2 - |z_2|^2) = w_1$. Hence, using the results in [20, Appendix 4.B], the pairwise error probability corresponding to $\mathcal{E}_0$ is

$$\begin{aligned} P(\mathcal{E}_0) &= P\{y > 0\} = P\{|z_1|^2 > |z_2|^2\} \\ &= Q_M\left(\frac{|\eta_{z_1}|}{\sqrt{2\sigma^2}}, \frac{|\eta_{z_2}|}{\sqrt{2\sigma^2}}\right) - \frac{1}{2} I_0\left(\frac{|\eta_{z_1}|\,|\eta_{z_2}|}{2\sigma^2}\right) \\ &\quad \cdot \exp\left(-\frac{|\eta_{z_1}|^2 + |\eta_{z_2}|^2}{4\sigma^2}\right) \end{aligned}$$

$$= \frac{1}{2} + \frac{1}{2} Q_M \left(\frac{|\eta_{z_1}|}{\sqrt{2\sigma^2}}, \frac{|\eta_{z_2}|}{\sqrt{2\sigma^2}} \right) - \frac{1}{2} Q_M \left(\frac{|\eta_{z_2}|}{\sqrt{2\sigma^2}}, \frac{|\eta_{z_1}|}{\sqrt{2\sigma^2}} \right) \quad (38)$$

where

$$I_0(x) \triangleq \frac{1}{2\pi} \int_0^{2\pi} \exp\{x \cos \alpha\} \, d\alpha \quad (39)$$

is the zeroth-order modified Bessel function of the first kind,

$$Q_M(x_1, x_2) \triangleq \int_{x_2}^{+\infty} \exp\left(-\frac{x_1^2 + x^2}{2} \right) I_0(x_1 x) x \, dx \quad (40)$$

is the Marcum Q function, and we have used the property [24]

$$Q_M(x_1, x_2) + Q_M(x_2, x_1) = 1 + \exp\left(-\frac{x_1^2 + x_2^2}{2} \right) I_0(x_1 x_2). \quad (41)$$

In (38), η_{z_1} and η_{z_2} are the sum of the conjugates of the elements of $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$, respectively,

$$\eta_{z_1} = (P-1) b^* \alpha^* + \alpha^*$$
$$\eta_{z_2} = -(P-1) b^* \alpha^* + \alpha^*. \quad (42)$$

The asymptotic bit-error probability for high SNR is

$$P_b \simeq \begin{cases} 2P(\mathcal{E}_0), & \text{for } M = 2 \quad \text{(BPSK)} \\ \dfrac{4}{\log_2 M} P(\mathcal{E}_0), & \text{for } M > 2 \quad (M\text{-PSK}). \end{cases} \quad (43)$$

In Fig. 7, this asymptotic bit-error probability for QPSK and various values of N is shown and compared with simulation results [12]. An excellent agreement with simulation results may be noted for low SNR as well, basically because less probable error events have negligible probability. In addition, this figure explicitly confirms that the performance approaches that of coherent detection for increasing values of N. In Section IV, the minimum equivalent distance for differentially encoded QPSK with $N = 5$ is found to be $d_{\text{eq}} = 1.33333$. This value must be compared to the Euclidean distance $d_E = \sqrt{2}$ which characterizes the coherent detection performance. This difference exactly matches the asymptotic high-SNR loss of 0.5 dB shown in Fig. 7, for $N = 5$.

Using this approach, we may explicitly verify, for differential encoding, that the high-SNR performance of these noncoherent detection schemes approaches that of coherent detection for increasing implicit phase memory. As mentioned in the Introduction, this result was previously obtained in [6] for asymptotically increasing block length and differentially encoded M-PSK. This result has been generalized to any coded PSK in Theorem 1. We now rederive this result on the basis of (43) by an alternative more explicit method which is also used in Appendix C.

For $N \to +\infty$, we have $P \to +\infty$, $|\eta_{z_1}| \to +\infty$, and $|\eta_{z_2}| \to +\infty$. Nevertheless, the difference $|\eta_{z_2}| - |\eta_{z_1}|$ remains finite. In fact, substituting in (42) the expression of a and b, letting $(a/|a|) = e^{j\beta}$, and recalling that $|\alpha| = (1/\sqrt{2})$, we have

$$|\eta_{z_1}| = \frac{1}{\sqrt{2}} \sqrt{P + 2\sqrt{P-1} \cos \beta}$$
$$|\eta_{z_2}| = \frac{1}{\sqrt{2}} \sqrt{P - 2\sqrt{P-1} \cos \beta} \quad (44)$$

and

$$\frac{|\eta_{z_2}| - |\eta_{z_1}|}{\sqrt{2\sigma^2}} = -\frac{1}{2\sigma} \frac{4\sqrt{P-1} \cos \beta}{\sqrt{P + 2\sqrt{P-1} \cos \beta} + \sqrt{P - 2\sqrt{P-1} \cos \beta}}. \quad (45)$$

Taking the limit as $P \to \infty$, and recalling that $\sigma^2 = (N_0/2E_S)$ and $\cos \beta = -\sin(\pi/M)$, we have

$$\lim_{P \to \infty} \frac{|\eta_{z_2}| - |\eta_{z_1}|}{\sqrt{2\sigma^2}} = \sqrt{\frac{2E_S}{N_0}} \sin \frac{\pi}{M}. \quad (46)$$

In this situation ($|\eta_{z_1}| \to +\infty$, $|\eta_{z_2}| \to +\infty$, $|\eta_{z_2}| - |\eta_{z_1}|$ finite), the following approximation holds [24]:

$$1 + Q_M(x_1, x_2) - Q_M(x_2, x_1) \simeq 2Q(x_2 - x_1). \quad (47)$$

Hence, for $N \to \infty$, (43) becomes

$$P_b \simeq \begin{cases} 2Q\left(\sqrt{\dfrac{2E_S}{N_0}} \right), & \text{for } M = 2 \text{ (BPSK)} \\ \dfrac{4}{\log_2 M} Q\left(\sqrt{\dfrac{2E_S}{N_0}} \sin \dfrac{\pi}{M} \right), & \text{for } M > 2 \ (M\text{-PSK}). \end{cases} \quad (48)$$

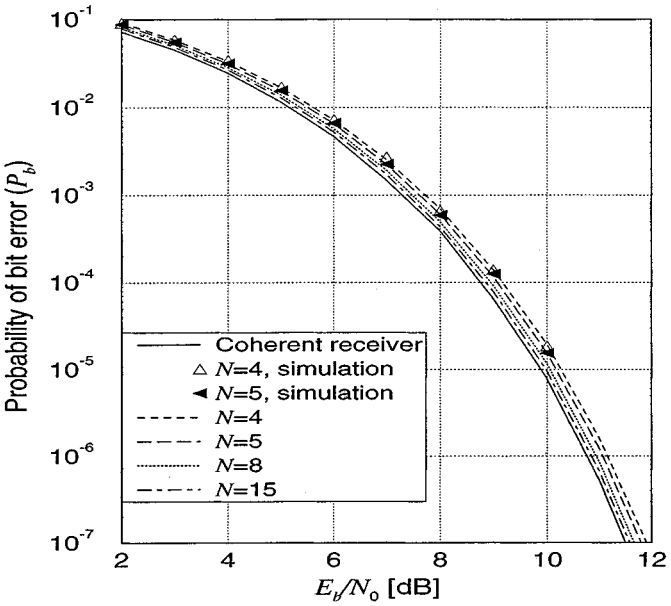

Fig. 7. Asymptotic performance for differentially encoded QPSK, various values of N, and comparison with simulation results.

This result may be recognized as the bit-error probability for coherent detection of differentially encoded M-PSK signals and, as expected, coincides with the asymptotic result given in [6].

VI. Conclusion

A theoretical performance analysis of new noncoherent combined sequence detection and decoding schemes for coded M-PSK [12] has been presented. Starting from the classical union bound, a method to numerically evaluate the pairwise error probability has been proposed and an upper bound on this probability derived. Based on this bound, it is possible to define an equivalent distance which plays a role similar to that of the Euclidean distance in coherent decoding and is useful in the performance analysis. An efficient search for optimal codes may be based on the minimum equivalent distance. This equivalent distance allows us to prove, in the general case of coded PSK, that the performance of the considered noncoherent schemes approaches that of coherent schemes when the phase memory parameter increases. The theoretical analysis has been shown to be in excellent agreement with computer simulation results. A simple expression of the asymptotic high-SNR bit-error probability of differentially encoded M-PSK signals has been obtained, which is in very good agreement with simulations for low SNR as well.

Appendix A

If P is odd, the determinant $|\boldsymbol{A}|$ of matrix $\boldsymbol{A}$ defined in (10) is zero. In fact, if we consider matrix $\boldsymbol{B}$ with $B_{ij}=\hat{c}_i c_j-\hat{c}_j c_i$, obtained from matrix $\boldsymbol{A}$ multiplying its first column by $\hat{c}_0 c_0$, its second by $\hat{c}_1 c_1$, and so on, we have

$$|\boldsymbol{B}|=\hat{c}_0 c_0 \hat{c}_1 c_1 \cdots \hat{c}_{P-1} c_{P-1}|\boldsymbol{A}|. \tag{A.1}$$

Since $\hat{c}_i, c_i \neq 0$ ($\hat{c}_i$ and c_i belong to the M-PSK alphabet), $|\boldsymbol{A}|=0$ if and only if $|\boldsymbol{B}|=0$. Matrix $\boldsymbol{B}$ is such that

$$B_{ji}=\hat{c}_j c_i-c_j\hat{c}_i=-(\hat{c}_i c_j-\hat{c}_j c_i)=-B_{ij}$$

i.e., $\boldsymbol{B}=-\boldsymbol{B}^T$. Using the property $|\boldsymbol{B}^T|=|\boldsymbol{B}|$, we have $|\boldsymbol{B}|=(-1)^P|\boldsymbol{B}|$. If P is odd this equality is verified only if $|\boldsymbol{B}|=0$.

In the characteristic polynomial of a matrix

$$p(\mu)=(-1)^P\mu^P+(-1)^{P-1}p_1\mu^{P-1}+\cdots+p_P \tag{A.2}$$

the coefficient p_i is the sum of the determinants of the principal minors of order i. Since these principal minors have the same structure of matrix $\boldsymbol{A}$, their determinants are zero if i is odd. Therefore, the characteristic polynomial includes only the powers $\mu^P, \mu^{P-2}, \mu^{P-4}, \cdots$. Moreover, its roots are real because $\boldsymbol{A}$ is Hermitian. It is straightforward to show that the nonzero roots of a polynomial with these characteristics are in equal and opposite pairs.

Appendix B

In this appendix, we derive (33). We assume that the transmitted code symbols are $c_k=\alpha_k+j\beta_k$, an error event starts at time k ($k=N-1$ for differential encoding and $k=N/\eta-1$ for convolutional encoding) and is characterized by at most D' errored code symbols $\hat{c}_k=\hat{\alpha}_k+j\hat{\beta}_k$. It is well known that the Euclidean distance of this error event is

$$\begin{aligned} d_E^2&=\sum_{i=k}^{k+D'-1}|\hat{c}_i-c_i|^2=2\sum_{i=k}^{k+D'-1}(1-\hat{\alpha}_i\alpha_i-\hat{\beta}_i\beta_i)\\ &=2\sum_{i=k}^{k+D'-1}\mathrm{Re}[(1-\hat{c}_i c_i^*)] \end{aligned} \tag{B.1}$$

having exploited the property $|c_k|^2=\alpha_k^2+\beta_k^2=1$. From (29) and (10), we have

$$\begin{aligned} d_{\mathrm{eq}}^2&=\frac{\left[\sum_{i=0}^{P-1}\sum_{j=0}^{P-1}c_i^*(\hat{c}_i\hat{c}_j^*-c_i c_j^*)c_j\right]^2}{\sum_{i=0}^{P-1}\sum_{j=0}^{P-1}\sum_{l=0}^{P-1}c_i^*(\hat{c}_i\hat{c}_l^*-c_i c_l^*)(\hat{c}_l\hat{c}_j^*-c_l c_j^*)c_j}\\ &=\frac{\left[\sum_{i=0}^{P-1}\sum_{j=0}^{P-1}(c_i^*c_j\hat{c}_i\hat{c}_j^*-1)\right]^2}{P\sum_{i=0}^{P-1}\sum_{j=0}^{P-1}(1-c_i^*c_j\hat{c}_i\hat{c}_j^*)}\\ &=\frac{\sum_{i=0}^{P-1}\sum_{j=0}^{P-1}(1-c_i^*c_j\hat{c}_i\hat{c}_j^*)}{P}. \end{aligned} \tag{B.2}$$

The right-hand side of (B.2) equals $|\eta_y|/P$. Recalling the definition of d_{eq} given by (28) we may conclude that

$$\frac{|\eta_y|}{s_y^2}=\frac{1}{4\sigma^2 P}.$$

Due to the fact that the maximum eigenvalue of $\boldsymbol{A}$ increases with $\sqrt{P}$ (see Appendix C), the condition

$$\zeta_0=\frac{|\eta_y|}{s_y^2}=\frac{1}{4\sigma^2P}\ll\frac{1}{2\mu_M\sigma^2} \tag{B.3}$$

is certainly verified when $N\to\infty$ and the approximate Chernoff bound (27) is asymptotically correct.

Equation (B.2) may be expressed as

$$\begin{aligned} d_{\mathrm{eq}}^2&=\left(1-\frac{D'}{P}\right)\sum_{i=k}^{k+D'-1}2\,\mathrm{Re}[(1-\hat{c}_i c_i^*)]\\ &\quad+\frac{2}{P}\sum_{i=k}^{k+D'-1}\sum_{j=k}^{i-1}\mathrm{Re}[(1-c_i^*c_j\hat{c}_i\hat{c}_j^*)]\\ &=\left(1-\frac{D'}{P}\right)d_E^2+\frac{1}{P}d_x^2 \end{aligned} \tag{B.4}$$

in which

$$d_x^2\triangleq 2\sum_{i=k}^{k+D'-1}\sum_{j=k}^{i-1}\mathrm{Re}[(1-c_i^*c_j\hat{c}_i\hat{c}_j^*)]. \tag{B.5}$$

$$\begin{cases} c_j \displaystyle\sum_{i=k}^{k+D'-1} (\hat{c}_i^* - c_i^*) q_i = \mu q_j, & \text{for } j = 0, 1, \cdots, k-1, k+D', \cdots, P-1 \\ (\hat{c}_j - c_j) \left[\displaystyle\sum_{i=0}^{k-1} c_i^* q_i + \sum_{i=k+D'}^{P-1} c_i^* q_i \right] + \sum_{i=k}^{k+D'-1} (\hat{c}_j \hat{c}_i^* - c_j c_i^*) q_i = \mu q_j, & \text{for } j = k, k+1, \cdots, k+D'-1. \end{cases} \tag{C.1}$$

$$\begin{cases} c_j \displaystyle\sum_{i=k}^{k+D'-1} (\hat{c}_i^* - c_i^*) q_i = \mu q_j, & \text{for } j = 0, 1, \cdots, k-1, k+D', \cdots, P-1 \\ (\hat{c}_j - c_j) \left[\displaystyle\sum_{i=0}^{k-1} c_i^* q_i + \sum_{i=k+D'}^{P-1} c_i^* q_i \right] = \mu q_j, & \text{for } j = k, k+1, \cdots, k+D'-1. \end{cases} \tag{C.3}$$

$$q_j = \begin{cases} \pm \dfrac{c_j \alpha}{\sqrt{P - D'}}, & \text{for } j = 0, 1, \cdots, k-1, k+D', \cdots, P-1 \\ \dfrac{(\hat{c}_j - c_j)\alpha}{d_E}, & \text{for } j = k, k+1, \cdots, k+D'-1 \end{cases} \tag{C.5}$$

APPENDIX C

In this appendix, we prove an asymptotic property of the eigenvalues of matrix $\boldsymbol{A}$ for $N \to \infty$. This result provides an alternative proof that, in the considered noncoherent schemes, the pairwise error probability of a generic error event tends to the corresponding pairwise error probability in coherent schemes.

Writing explicitly the equation $\boldsymbol{Aq} = \mu \boldsymbol{q}$ in order to find the eigenvalues and eigenvectors of $\boldsymbol{A}$, we obtain (C.1), shown at the top of this page. If $P \gg D'$ and the condition

$$\sum_{i=k}^{k+D'-1} (\hat{c}_j \hat{c}_i^* - c_j c_i^*) q_i \ll (\hat{c}_j - c_j) \left[\sum_{i=0}^{k-1} c_i^* q_i + \sum_{i=k+D'}^{P-1} c_i^* q_i \right] \tag{C.2}$$

is verified for any value of j such that $k \geq j \geq k + D' - 1$, (C.1) becomes (C.3), shown at the top of this page. The solutions of (C.3) are

$$\mu_{1,2} = \pm \sqrt{P - D'}\, d_E \tag{C.4}$$

and (C.5), also at the top of this page, where $|\alpha|^2 = \frac{1}{2}$. With these solutions, condition (C.2) becomes

$$\frac{\displaystyle\sum_{i=k}^{k+D'-1} (\hat{c}_j \hat{c}_i^* - c_j c_i^*)}{d_E} \ll \sqrt{P - D'} \tag{C.6}$$

which is verified when P is large enough. We may conclude that, when $P \to \infty$, matrix $\boldsymbol{A}$ has only two eigenvalues and the corresponding eigenvectors are given by (C.5). Following the derivation in Section V, the pairwise error probability satisfies

$$\lim_{N \to \infty} P(\boldsymbol{a} \to \hat{\boldsymbol{a}}) = Q\left(\frac{d_E}{2\sigma} \right). \tag{C.7}$$

REFERENCES

[1] P. Y. Kam, "Maximum-likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.*, vol. COM-34, pp. 522–527, June 1986.

[2] H. Leib and S. Pasupathy, "The phase of a vector perturbed by Gaussian noise and differentially coherent receivers," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1491–1501, Nov. 1988.

[3] F. Edbauer, "Bit error rate of binary and quaternary DPSK signals with multiple differential feedback detection," *IEEE Trans. Commun.*, vol. 40, pp. 457–460, Mar. 1992.

[4] P. Y. Kam, S. S. Ng, and T. S. Ng, "Optimum symbol-by symbol detection of uncoded digital data over the Gaussian channel with unknown carrier phase," *IEEE Trans. Commun.*, vol. 42, pp. 2543–2551, Aug. 1994.

[5] S. G. Wilson, J. Freebersyser, and C. Marshall, "Multi-symbol detection of M-DPSK," in *Proc. Global Communications Conf. (GLOBECOM'89)*, Nov. 1989, pp. 1692–1697.

[6] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, pp. 300–308, Mar. 1990.

[7] H. Leib and S. Pasupathy, "Noncoherent block demodulation of PSK," in *Proc. Vehicular Technology Conf. (VTC'90)*, May 1990, pp. 407–411.

[8] D. Divsalar, M. K. Simon, and M. Shahshahani, "The performance of trellis-coded MDPSK with multiple symbol detection," *IEEE Trans. Commun.*, vol. 38, pp. 1391–1403, Sept. 1990.

[9] D. Raphaeli, "Noncoherent coded modulation," *IEEE Trans. Commun.*, vol. 44, pp. 172–183, Feb. 1996.

[10] Y. Kofman, E. Zehavi, and S. Shamai (Shitz), "nd-convolutional codes-part I: Performance analysis," *IEEE Trans. Inform. Theory*, vol. 43, pp. 558–575, Mar. 1997.

[11] ——, "nd-convolutional codes—Part II: Structural analysis," *IEEE Trans. Inform. Theory*, vol. 43, pp. 576–589, Mar. 1997.

[12] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, pp. 1376–1385, Sept. 1999.

[13] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 74–83, Jan. 1992.

[14] C. W. Helstrom, *Elements of Signal Detection and Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[15] D. Raphaeli, "Distribution of noncentral indefinite quadratic forms in complex normal variables," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1002–1007, May 1996.

[16] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Simple method for evaluating error probabilities," *Electron. Lett.*, vol. 32, no. 3, pp. 191–192, Feb. 1996.

[17] T. Aulin and C.-E. Sundberg, "Partially coherent detection of digital full response continuous phase modulated signals," *IEEE Trans. Commun.*, vol. COM-30, pp. 1096–1117, May 1982.

[18] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[19] D. Raphaeli, "On multidimensional coded modulations having uniform error property for generalized decoding and flat-fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 34–40, Jan. 1998.

[20] J. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[21] M. Schwartz, W. R. Bennet, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[22] G. Colavolpe. (1998, Feb.) "Ricevitori a rivelazione di sequenza per trasmissioni su canali a fase incognita," Ph.D. dissertation. Università di Parma, Parma, Italy. [Online] Available: http://com.unipr.it/people/colavolpe/giulio.html.

[23] V. K. Prabhu, "Modified Chernoff bounds for PAM systems with noise and interference," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 95–100, Jan. 1982.

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