

# On Linear Predictive Detection for Communications With Phase Noise and Frequency Offset

Gianluigi Ferrari, *Member, IEEE*, Giulio Colavolpe, *Member, IEEE*, and Riccardo Raheli, *Member, IEEE*

**Abstract**—In this paper, by applying the concept of linear prediction, which is widely used for fading channels, to phase-uncertain communications, we generalize existing linear predictive detection algorithms for transmission over channels with phase noise and frequency offset. This approach leads to the derivation of detection algorithms, which are referred to as phasor linear predictive (pLP), for trellis-based maximum *a posteriori* (MAP) sequence detection (based on the Viterbi algorithm) and MAP symbol detection: trellis-based (using the forward–backward algorithm) and graph-based (using the sum-product algorithm). The effectiveness of the proposed pLP detection algorithms is evaluated for several communication schemes. The derived algorithms outperform previously appeared finite-memory detection solutions in terms of robustness against fast channel dynamics. Moreover, the proposed detection strategy lends itself to attractive extensions to adaptive schemes.

**Index Terms**—Finite-memory detection, forward–backward (FB) algorithm, graph-based detection, iterative detection, linear prediction, maximum *a posteriori* (MAP) sequence/symbol detection, minimum mean square error (MMSE) estimation, sum-product (SP) algorithm, Viterbi algorithm (VA).

## I. INTRODUCTION

PHASE uncertainty, which is due to phase noise or frequency offset, is one of the most detrimental and common impairments of radio communication systems [1]. In addition to cellular networks and satellite communications [2], in future-generation wireless local area networks, the integration of mobile and satellite communications will call for detection algorithms suitable to phase-uncertain channels. In particular, these detection algorithms should be robust to oscillator instabilities, generating phase noise, and time-varying frequency offsets, possibly due to the Doppler shift and experienced, for example, in low Earth orbit (LEO) satellite systems [2], [3].

Various detection strategies have appeared in the literature to combat the effects of channel phase uncertainty. These

Manuscript received September 22, 2004; revised March 30, 2005, August 8, 2005, December 17, 2005, and June 5, 2006. This work was presented in part at the European Association for Signal Processing (EURASIP)/IEEE International Symposium on Signal Processing and Applications, July 2003, Paris, France. The review of this paper was coordinated by Dr. M. Stojanovic.

The authors are with the Department of Information Engineering, University of Parma, 43100 Parma, Italy (e-mail: gianluigi.ferrari@unipr.it; giulio@unipr.it; raheli@unipr.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2007.897190

strategies can be classified into three main categories as follows.

- 1) The first group of trellis-based detection algorithms is based on the key idea of per-path recursive computation of a channel phase estimate. This represents an application of the concept of per-survivor processing (PSP) [4]. In particular, in [4] and [5], hard-output maximum *a posteriori* (MAP) sequence detection algorithms, based on the Viterbi algorithm (VA) [6], [7], are derived. In [8], this approach is extended to iterative detection schemes based on the use of the forward–backward (FB) algorithm [9]. In particular, in [8], recursive least mean squares estimation of the unknown channel phase is considered in the computations that are involved in the proposed FB algorithm instances. Schober and Gerstacker [10] propose a noncoherent sequence detection (NSD) algorithm with per-path recursive metric computation based on a heuristic modification of the metric proposed in [11].
- 2) The second group of detection algorithms is based on the application of a general *finite-memory condition*. The basic idea consists in limiting *a priori* the channel memory. This approach was considered in [12] and [13]. In [11], an instance of this detection philosophy for VA-based MAP sequence detection algorithms, which are referred to as NSD, is proposed. In this case, the channel phase is assumed to be a random variable that is uniformly distributed in  $[0, 2\pi]$ . This approach is extended to the case of iterative detection, based on the FB algorithm, in [14] and [15]. A general framework for finite-memory detection is proposed in [16]. The class of detection algorithms proposed in this paper belongs to this group.
- 3) The third group of detection algorithms is obtained by modeling the channel phase process as a suitable Markov process. In [17], the channel phase process that is modeled as first-order Markov chain and detection algorithms, run over an expanded trellis diagram that takes into account the phase evolution, are derived. A general approach for maximum-likelihood detection over finite-state Markov channels is also considered in [18].

Linear prediction has been extensively considered for detection over *fading* channels [19]–[22]. In [23]–[25], linear prediction is used for noncoherent detection, over fading channels, with differential phase shift keying (DPSK). We point out that, in [23]–[25], only DPSK is considered, whereas the approach

presented in our paper is general and can be applied to any modulation format. In this paper, we exploit the concept of linear prediction to derive an optimal metric for finite-memory detection, and we show how to use this metric in various detection algorithms for *phase-uncertain* communications, which belong to the second group previously indicated. We will refer to this class of detection algorithms as phasor linear predictive (pLP) detection algorithms. Although stemming from a fundamentally different criterion for the estimation of the unknown channel phase process, the proposed pLP algorithms are related to previously appeared noncoherent detection schemes [11], [14], [26], exhibiting improved robustness to strong phase instabilities or uncompensated frequency offsets. In particular, it is worth pointing out that the proposed class of pLP detection algorithms represents a conceptual generalization, to time-varying phase-uncertain channels, of the noncoherent detection algorithm proposed in [11] and valid for channels introducing a random phase. This relation is investigated in detail in the Appendix.

We will first derive VA-based pLP detection algorithms for linear coded modulations, and we will then propose a possible extension, based on oversampling, to the case of continuous phase modulation (CPM) [27], [28]. The proposed class of VA-based pLP detection algorithms will be systematically extended to FB and sum-product (SP) [29] algorithms according to the framework in [16]. Numerical results will be presented for hard-output detection schemes (based on the VA) and soft-output iterative detection schemes (based on the FB and SP algorithms).

This paper is organized as follows. In Section II, pLP algorithms, both trellis-based and graph-based, are derived for a generic channel phase process, along with simplified versions and extensions to oversampled receivers. In Section III, the important case of a Wiener channel phase process with a time-invariant frequency offset is considered. Extensive numerical results relative to linear predictive detection schemes are presented in Section IV. In Section V, we comment on adaptive extensions of the proposed pLP detection algorithms to scenarios with time-varying channel dynamics. Section VI concludes the paper. In the Appendix, the case with a time-invariant channel phase is interpreted as a special instance of the general case considered in this paper.

## II. pLP DETECTION ALGORITHMS

We first derive VA-based pLP detection algorithms for linear coded modulations and then extend them to the case of CPM. A simplified version of VA-based pLP algorithms for nonequal energy signaling is also proposed. Finally, systematic extensions to pLP-FB and pLP-SP algorithms are outlined.

### A. Linear Coded Modulation

We consider the low-pass complex equivalent system depicted in Fig. 1. We assume that a sequence of  $K$  independent and uniformly distributed  $M$ -ary symbols  $\{a_k\}_{k=0}^{K-1}$ , which are also denoted by  $a$  for conciseness, feeds an encoder/modulator, which can be modeled as a time-invariant finite-state machine

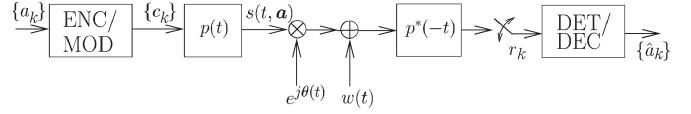


Fig. 1. System model.

(FSM) with state  $\mu_k$ . The evolution of the encoder/modulator FSM is described by a “next-state” function  $ns(\cdot, \cdot)$  [such that  $\mu_{k+1} = ns(\mu_k, a_k)$ ] and an “output” function  $o(\cdot, \cdot)$  [such that the output symbol<sup>1</sup> is  $c_k = o(\mu_k, a_k)$ ]. We also define the relevant state transition as  $t_k \stackrel{\Delta}{=} (a_k, \mu_k)$ . The linearly modulated continuous-time signal  $s(t, a)$  is obtained by letting the code symbol  $c_k$  be carried by a suitable shaping pulse  $p(t)$ . Although suboptimal in the presence of a time-varying channel, a matched-filter front-end with a sampling rate of one sample per symbol can be practically used, provided that the phase process is not affected by very strong variations [30]. The resulting observation model<sup>2</sup> is

$$r_k = c_k e^{j\theta_k} + w_k \quad (1)$$

where  $\{w_k\}$  is an independent identically distributed (i.i.d.) complex additive Gaussian noise sequence with variance  $N_0$ , and  $\{\theta_k\}$  is a zero-mean wide sense stationary (WSS) channel phase process. The autocorrelation sequence of the stationary phasor process  $\{e^{j\theta_k}\}$  is denoted by  $R_\theta(n) \stackrel{\Delta}{=} \mathbb{E}\{e^{j\theta_{k+n}} e^{-j\theta_k}\}$  and is assumed to be known at the receiver side. In the absence of this information at the receiver, the phasor autocorrelation function needs to be properly estimated. Obviously, the observables are not Gaussian, conditionally on the data sequence.

Should the phase process  $\{\theta_k\}$  be known exactly, a coherent sequence detector based on the VA, which was run over a trellis diagram relative to the encoder/modulator FSM, could be derived, with the branch metric

$$\begin{aligned} \lambda_k^{\text{coher}}(t_k) &\stackrel{\Delta}{=} -\frac{|r_k - e^{j\theta_k} c_k|^2}{N_0} + \ln P\{a_k\} \\ &= -\frac{|c_k|^2 |r'_k - e^{j\theta_k}|^2}{N_0} + \ln P\{a_k\} \end{aligned} \quad (2)$$

where  $r'_k \stackrel{\Delta}{=} r_k / c_k$  is a normalized observation—recall that a VA based on the metric (2) performs recursive maximization of path metrics associated with the survivors. By assuming that  $P\{a_k\} = 1/M$  (i.e., maximum likelihood detection), a metric equivalent to (2) is given as follows:

$$\lambda_k^{\text{coher}}(t_k) = -|r_k - e^{j\theta_k} c_k|^2 = -|c_k|^2 |r'_k - e^{j\theta_k}|^2. \quad (3)$$

<sup>1</sup>For simplicity, we refer to a scalar output function, but the extension to the case of multiple output symbols is straightforward, by the use of vector notation.

<sup>2</sup>For the sake of notational simplicity, we do not use a different formalism to distinguish between a random variable and its realization. The context should make it clear to which case the used notation refers.

The performance of a receiver using a VA with metric (2) [or (3)] will be considered as a benchmark for the performance of the proposed pLP-based sequence detection schemes.

If only a statistical characterization of the phase process is available, a practical sequence detector can be obtained by using (2) with the exact value  $\theta_k$  replaced by a suitable estimate  $\hat{\theta}_k$ . In the general case of a time-varying phase process, the data-aided minimum mean square error (MMSE) phase estimate based on  $\nu$  previous observations is given by the following conditional mean:

$$\hat{\theta}_k = \mathbb{E} \{ \theta_k | \mathbf{c}_{k-\nu}^{k-1}, \mathbf{r}_{k-\nu}^{k-1} \} \quad (4)$$

where an indexed vector notation is used to denote code symbols and observations from epoch  $k-\nu$  to epoch  $k-1$ . The expectation in (4) leads to a nonlinear estimate, which is usually rather difficult to compute.

Instead of directly estimating the phase, we use the following indirect estimation strategy. We note that for a large signal-to-noise ratio (SNR), one can write  $r'_k \simeq e^{j\theta_k}$  so that  $\theta_k \simeq \arg\{r'_k\}$ . In order to exploit the phase correlation characteristics in the estimation process, we consider a data-aided linear prediction  $\hat{r}'_k$  of  $r'_k$  based on the previous  $\nu$  normalized observations  $\{r'_j\}_{j=k-\nu}^{k-1}$ , i.e., we define

$$\hat{r}'_k \triangleq \sum_{i=1}^{\nu} p_i r'_{k-i} \quad (5)$$

where  $\{p_i\}_{i=1}^{\nu}$  are the *prediction coefficients*, and  $\nu$  is the *prediction order* [31]. The prediction coefficients in (5) can be computed by solving the following MMSE problem [19]–[22]:

$$\min_{\{p_i\}_{i=1}^{\nu}} \mathbb{E} \left\{ \left| r'_k - \sum_{i=1}^{\nu} p_i r'_{k-i} \right|^2 \middle| \mathbf{c}_{k-\nu}^{k-1} \right\} \quad (6)$$

which leads to a Wiener–Hopf linear system  $\mathbf{R}\mathbf{p} = \mathbf{b}$ , where  $\mathbf{R}$  is a square  $\nu \times \nu$  matrix whose elements have the following expression:

$$[\mathbf{R}]_{l,m} = \begin{cases} R_\theta(|l-m|), & \text{if } l \neq m \\ R_\theta(0) + \frac{N_0}{|c_{k-l}|^2}, & \text{if } l = m \end{cases} \quad (7)$$

where  $\mathbf{p} \triangleq [p_1 \cdots p_{\nu}]^T$  is the unknown vector, and  $\mathbf{b} \triangleq [R_\theta(1), R_\theta(2), \dots, R_\theta(\nu)]^T$ . With the prediction coefficients obtained by solving the Wiener–Hopf system, it is easy to express the corresponding MMSE, which is denoted as  $\epsilon_k^2$ , as follows:

$$\epsilon_k^2 = R_\theta(0) + \frac{N_0}{|c_k|^2} - \sum_{i=1}^{\nu} p_i R_\theta(i). \quad (8)$$

One can therefore consider the following indirect data-aided phase estimate:

$$\hat{\theta}_k \triangleq \arg \left\{ \hat{r}'_k \right\} = \arg \left\{ \sum_{i=1}^{\nu} p_i r'_{k-i} \right\} = \arg \left\{ \sum_{i=1}^{\nu} p_i \frac{r_{k-i}}{c_{k-i}} \right\}. \quad (9)$$

Upon the definition of an extended state<sup>3</sup>  $S_k \triangleq (a_{k-1}, \dots, a_{k-\nu}, \mu_{k-\nu})$ , it is immediate to conclude that the optimal prediction coefficients depend, in general, on  $S_k$  and that the MMSE depends on the extended transition  $T_k \triangleq (a_k, S_k)$ —note that an extended transition  $T_k$  uniquely identifies a sequence of code symbols  $c_{k-\nu}^k$ . Defining an extended trellis diagram (with respect to the trellis diagram relative to the encoder/modulator FSM) with state  $S_k$ , a VA with the following branch metric can be implemented:

$$\begin{aligned} \lambda_k^{\text{pLP}}(T_k) &= -\frac{|c_k|^2 \left| r'_k - e^{j\hat{\theta}_k} \right|^2}{\epsilon_k^2} + \ln P\{a_k\} \\ &= -\frac{|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^{\nu} p_i r'_{k-i}}{\left| \sum_{i=1}^{\nu} p_i r'_{k-i} \right|} \right|^2}{\epsilon_k^2} + \ln P\{a_k\}. \end{aligned} \quad (10)$$

A VA based on metric (10) can be interpreted as performing combined detection and decoding, with per-survivor linear predictive estimation based on  $\nu$  consecutive observations, of the channel phase process.

In the case of equal energy signaling, since the system matrix  $\mathbf{R}$  in (7) no longer depends on  $\{c_{k-l}\}_{l=1}^{\nu}$ , the prediction coefficients and the MMSE do not depend on  $T_k$ , but only on the SNR. In this case, they can be precomputed offline and then used in the VA. Moreover, if a symbol can assume with equal probability all possible values, the metric in (10) can be written as follows:

$$\lambda_k^{\text{pLP}}(T_k) = -|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^{\nu} p_i r'_{k-i}}{\left| \sum_{i=1}^{\nu} p_i r'_{k-i} \right|} \right|^2. \quad (11)$$

With respect to classical linear predictive receivers for fading channels [19]–[22], where the fading amplitude is not unitary, the proposed solution features the main difference, i.e., the fact that the phasor estimate has (obviously) unit modulus. In particular, this is obtained by normalizing  $\sum_{i=1}^{\nu} p_i r'_{k-i}$  by its modulus. Our results show that the absence of this normalization, although not rigorous, entails a minor performance degradation for equal-energy modulations. However, this normalization is essential for nonequal energy signaling such as quadrature amplitude modulations (QAMs), as will be shown in Section IV.

The complexity of the proposed VA-based pLP detection algorithms can be limited by applying reduced-state sequence detection (RSSD) techniques [33]–[35]. As an example, a reduced state can be defined by substituting  $\nu$  with a reduced-state parameter  $Q < \nu$  in the definition of  $S_k$ , i.e., by “memory truncation.” In other words, the reduced state is defined as  $s_k \triangleq (a_{k-1}, \dots, a_{k-Q}, \mu_{k-Q})$ . During the recursive computation of the VA, decision feedback sequences that are associated

<sup>3</sup>The idea of expanding the trellis diagram to partially take into account the channel memory can be originally traced back to carrier phase recovery for analog frequency-modulated communications [32].

with each trellis state are recursively updated and used in the computation of the branch metrics. In the case of high-order modulation formats, other state reduction techniques, based on set partitioning [33], [34], can be used.

We remark that the prediction coefficients  $\{p_i\}_{i=1}^\nu$  can also be computed by an adaptive algorithm that recursively minimizes the mean square error in (6), such as a stochastic gradient algorithm [36]. This suggests that, on the practical side, the proposed algorithms lend themselves to the implementation, through standard methods, of adaptive receivers that are capable of coping with possible time variations of the statistics of the underlying channel phase model. This aspect will be addressed in more detail in Section V.

### B. Simplified pLP Detection Algorithms for Nonequal Energy Linear Modulations

Considering the general Wiener–Hopf system with a matrix given by (7), it is immediate to conclude that the prediction coefficients to be used for the computation of the branch metric  $\lambda_k^{\text{pLP}}(T_k)$  depend, in general, on the transition  $T_k$ . This implies that the ensemble of all possible transitions determines an ensemble of prediction coefficient sets. More precisely, by indicating by  $M_{\text{amp}}$  the number of different symbol *amplitudes* for the constellation relative to the nonequal energy modulation format, it can easily be concluded that  $M_{\text{amp}}^N$  sets of  $\nu$  prediction coefficients are required in the detection process. In this subsection, we propose a simple approximation to derive a unique set of prediction coefficients, independent from the specific transition, also in the case of nonequal energy modulations. In Section IV, it will be shown that the performance degradation entailed by this simplification is not significant.

Consider a modified MMSE problem, where the squared error is averaged with respect to both channel phase and code symbols. In other words, the prediction coefficients are obtained by solving the following minimization problem:

$$\begin{aligned} \min_{\{p_i\}_{i=1}^\nu} \mathbb{E} & \left\{ \left| r'_k - \sum_{i=1}^\nu p'_i r'_{k-i} \right|^2 \right\} \\ & = \min_{\{p_i\}_{i=1}^\nu} \mathbb{E} \left\{ \mathbb{E} \left\{ \left| r'_k - \sum_{i=1}^\nu p'_i r'_{k-i} \right|^2 \middle| \mathbf{c}_{k-\nu}^{k-1} \right\} \right\}. \quad (12) \end{aligned}$$

The minimization problem in (12) leads to a Wiener–Hopf system of the form  $\mathbf{R}_{\text{simp}} \mathbf{p}' = \mathbf{b}$ , where a generic element  $[\mathbf{R}_{\text{simp}}]_{i,j}$  of the system matrix  $\mathbf{R}_{\text{simp}}$  has the following expression:

$$[\mathbf{R}_{\text{simp}}]_{i,j} = \begin{cases} R_\theta(|i-j|), & \text{if } i \neq j \\ R_\theta(0) + N_0 \mathbb{E} \left\{ \frac{1}{|c_{k-i}|^2} \right\}, & \text{if } i = j. \end{cases} \quad (13)$$

Assume that a code symbol  $c_k$  can be any point of the constellation with equal probability: It follows that the quantity  $\mathbb{E}\{1/|c_{k-i}|^2\}$  depends only on the specific constellation. For

example, in the case of 16-ary QAM (16-QAM), it is easy to show that

$$\mathbb{E} \left\{ \frac{1}{|c_{k-i}|^2} \right\} \simeq \frac{1.89}{E_S} \quad (14)$$

where  $E_S$  is the average symbol energy.

### C. CPM

The decomposition approach to CPM proposed in [37] clearly shows that any CPM modulator can be interpreted as a serial concatenation of a continuous phase encoder (CPE) and a memoryless mapper. In particular, the CPE is a rate-1/ $L$  recursive encoder with state  $\mu_k = (a_{k-1}, \dots, a_{k-L+1}, \rho_{k-L+1})$ , where  $L$  is the duration of the frequency impulse that characterizes the CPM signal, and  $\rho_k$  is a  $p$ -ary symbol, which is recursively updated according to  $\rho_k = R_p[\rho_{k-1} + a_k]$ , where  $R_p[x]$  indicates the remainder of  $x$  in base  $p$ . The integer  $p$  is related to the CPM modulation index  $h$ : In fact,  $h = k/p$ , with  $k$  and  $p$  relatively prime numbers. The equivalent baseband CPM signal  $s_{\text{bb}}(t, \mathbf{a})$  can be expressed as [37] follows:

$$s(t, \mathbf{a}) = \exp\{j\bar{\psi}(t, \mathbf{a})\} \quad (15)$$

in which the *tilted phase*  $\bar{\psi}$  has the following expression:

$$\begin{aligned} \bar{\psi}(\tau + kT, \mathbf{a}) & = R_{2\pi} \left[ 2\pi h \rho_{k-L+1} \right. \\ & \quad \left. + 4\pi h \sum_{i=0}^{L-1} a_{k-i} q(\tau + iT) + \xi(\tau) \right], \quad 0 \leq \tau \leq T \quad (16) \end{aligned}$$

where  $q(t)$  is the CPM phase pulse, and  $\xi(\tau)$  is a function of time only (i.e., it does not depend on any information symbol) [37].

The receiver front-end can be implemented, in a practically optimal way, with the use of a simple filter followed by oversampling [38]. Considering  $\beta$  samples per symbol interval, the generic sample in the  $k$ th interval has the following expression:

$$\begin{aligned} r_k^{(i)} & = e^{j\theta_k^{(i)}} s_k^{(i)} + n_k^{(i)} \\ & = \exp\left\{j\left[\theta_k^{(i)} + \bar{\psi}(iT/\beta + kT, \mathbf{a})\right]\right\} + n_k^{(i)} \\ & \quad i = 0, \dots, \beta - 1 \quad (17) \end{aligned}$$

where  $s_k^{(i)} \triangleq s(iT/\beta + kT, \mathbf{a})$ . The bandwidth expansion (due to oversampling) in the front-end bandpass filter makes the variance of the Gaussian noise sample  $\{n_k^{(i)}\}$  equal to  $\beta N_0$ . In the case of an AWGN channel, being the additive channel noise white, it follows that  $\{n_k^{(i)}\}$  may be made independent by suitably setting the filter frequency response.

As a possible direct extension of the linear predictive strategy proposed in Section II-A, and based on a sampling rate of one sample per symbol interval, we now assume that  $e^{j\theta_k^{(i)}}$  is estimated based on the observation sequence  $(r_{k-N}^{(i)'}, \dots, r_{k-N}^{(\beta-1)'}, \dots, r_{k-1}^{(0)'}, \dots, r_{k-1}^{(\beta-1)'}, r_k^{(0)'}, \dots, r_k^{(i-1)'})$ . Fig. 2 highlights the sliding window prediction strategy considered in the

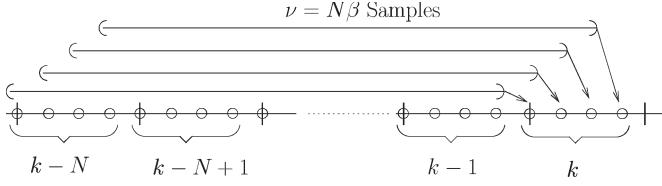


Fig. 2. Sliding window linear predictive detection strategy in the case of oversampling with  $\beta = 4$ .

case of oversampling. It is worth observing that different strategies based on observation windows of different lengths could be considered.<sup>4</sup> According to the current estimation strategy, the prediction order can be written as  $\nu = N\beta$ , and the basic metric, which is indicated by  $\lambda_k^{\text{PLP-o}}$ , becomes

$$\lambda_k^{\text{PLP-o}}(T_k) = -\frac{\sum_{i=0}^{\beta-1} |c_k^{(i)}|^2 \left| r_k^{(i)\prime} - \widehat{e^{\theta_k^{(i)}}} \right|^2}{(\epsilon_k^o)^2} + \ln P\{a_k\} \quad (18)$$

where  $(\epsilon_k^o)^2$  is the MMSE in the current case with oversampling. It is possible to extend the MMSE problem in (6) as follows:

$$\min_{\{p_i^o\}_{i=1}^\nu} \mathbb{E} \left\{ \left| r_k^{(i)\prime} - \sum_{j=0}^{i-1} p_{i-j}^o r_{k-i}^{(j)\prime} - \sum_{l=0}^{N-2} \sum_{j=0}^{\beta-1} p_{l+\beta+j}^o r_{k-1-l}^{(j)\prime} - \sum_{j=i}^{\beta-1} p_{N\beta+i-j}^o r_{k-N}^{(j)\prime} \right|^2 \middle| \mu_k \right\} \quad (19)$$

where  $\{p_i^o\}$  represents the prediction coefficients for the considered oversampled detection strategy. Solving the MMSE problem (19) for the prediction coefficients  $\{p_i^o\}$  leads to a Wiener–Hopf system of the form  $\mathbf{R}^o p^o = b^o$ , where  $\mathbf{R}^o$  is a square  $\nu \times \nu$  matrix whose elements, owing to the fact that  $|c_k^{(i)}| = 1$ , have the following expression:

$$[\mathbf{R}^o]_{l,m} = \begin{cases} R_\theta^{o,\beta}(|l-m|), & \text{if } l \neq m \\ R_\theta^{o,\beta}(0) + N_0, & \text{if } l = m \end{cases} \quad (20)$$

where  $\{R_\theta^{o,\beta}(j)\}$  is the autocorrelation sequence of the oversampled phase process  $\{\theta_k^{(i)}\}$ . Finally, the MMSE in the case of oversampling is formally identical to the case without oversampling, i.e.,

$$(\epsilon_k^o)^2 = R_\theta^{o,\beta}(0) + N_0 - \sum_{i=1}^\nu p_i^o R_\theta^{o,\beta}(i). \quad (21)$$

Note that being CPM a particular equal-energy modulation format, the prediction coefficients  $p^o$  and the MMSE  $(\epsilon_k^o)^2$  do not depend on the code symbols, i.e., neither on state  $S_k$  nor on transition  $T_k$ .

<sup>4</sup>For instance, in order to estimate  $e^{j\theta_k^{(i)}}$ , one could use all past observations down to  $r_{k-N}^{(0)\prime}$ . In this case, Fig. 2 should be modified in such a way that the prediction window for each sample in the  $k$ th symbol interval started from the initial sample of the  $(k-N)$ th symbol interval.

#### D. FB Algorithm: Trellis-Based MAP Symbol Detection

The derivation of a VA in the preceding subsections allows one to perform hard-output MAP *sequence* detection. According to the general framework on finite-memory detection in [16], the basic “exponential” metric to be used in a pLP-FB algorithm is given as follows:

$$\gamma_k^{\text{PLP}}(T_k) \triangleq \exp \left[ \lambda_k^{\text{PLP}}(T_k) \right] = \exp \left( -\frac{|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^\nu p_i r'_{k-i}}{\left| \sum_{i=1}^\nu p_i r'_{k-i} \right|} \right|^2}{\epsilon_k^2} \right) P\{a_k\} \quad (22)$$

and the *a posteriori* probability  $P\{a_k | \mathbf{r}\}$  can be approximately computed as

$$P\{a_k | \mathbf{r}\} \simeq \sum_{T_k : a_k} \alpha_k(S_k) \gamma_k^{\text{PLP}}(T_k) \beta_{k+1}(S_{k+1}) \quad (23)$$

where the notation  $T_k : a_k$  indicates all transitions  $T_k$  compatible with  $a_k$ . The quantities  $\alpha_k(S_k)$  and  $\beta_{k+1}(S_{k+1})$  can be computed by means of the following forward and backward recursions, respectively:

$$\alpha_{k+1}(S_{k+1}) = \sum_{T_k : S_{k+1}} \alpha_k(S_k) \gamma_k^{\text{PLP}}(T_k) \quad (24)$$

$$\beta_k(S_k) = \sum_{T_k : S_k} \beta_k(S_{k+1}) \gamma_k^{\text{PLP}}(T_k). \quad (25)$$

Proper boundary conditions for  $\alpha_0(S_0)$  and  $\beta_K(S_K)$  have to be considered. For example, considering a sequence of  $\nu$  pilot symbols at the beginning of the transmission, it is possible to set  $\alpha_0(S_0) = 1$  if  $S_0 = 0$ , and zero in correspondence of all other states.

The complexity of the proposed pLP-FB algorithms can be limited by applying the state reduction techniques proposed in [39]–[41]. In this case as well, a reduced-state trellis diagram is obtained by replacing  $\nu$  with a reduced-state parameter  $Q$  and applying decision feedback in at least one of the two recursions.

#### E. SP Algorithm: Graph-Based MAP Symbol Detection

As described in [16] and [42], the basic “exponential” metric (22) characterizing a pLP-FB algorithm can be associated with a factor node modeling the channel and connected to the factor graph describing the code constraints [29]. Given a code, it is possible to associate to it an indicator function  $\chi(c)$  such that

$$\chi(c) = \begin{cases} 1, & \text{if } c \text{ is a codeword} \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

The joint *a posteriori* probability of the code symbols, given the received observation sequence and discarding irrelevant terms,

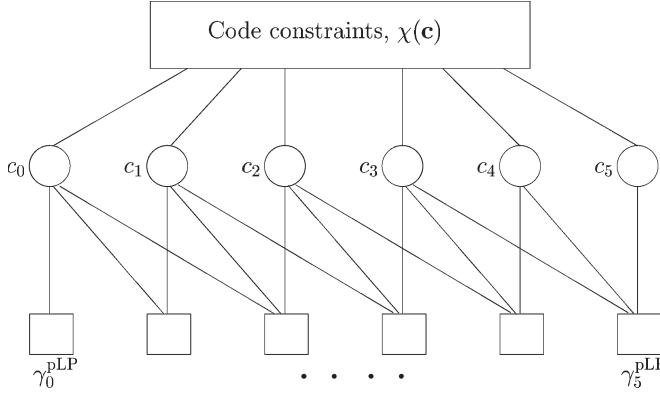


Fig. 3. Factor graph corresponding to the factorization (27) for  $\nu = 2$ .

can be approximately expressed as [42]

$$P(\mathbf{c}|\mathbf{r}) \simeq \chi(\mathbf{c}) \prod_k \gamma_k^{\text{PLP}} (\mathbf{c}_{k-\nu}^k) \quad (27)$$

where

$$\gamma_k^{\text{PLP}} (\mathbf{c}_{k-\nu}^k) \triangleq \exp \left( -\frac{|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^{\nu} p_i r'_{k-i}}{\sum_{i=1}^{\nu} p_i} \right|^2}{\epsilon_k^2} \right) P\{a_k\}. \quad (28)$$

Comparing (28) with (22), it can easily be recognized that they are identical, but for a formal difference. Although the FB algorithm is trellis based and, consequently, suitable definitions of state and transition are considered, in the case of graph-based detection on a factor graph, the definitions of state and transition become irrelevant.

The factor graph relative to the factorization in (27) is pictured in Fig. 3. As one can see, this factor graph takes into account both the code constraints [described by  $\chi(\mathbf{c})$ ] and the channel statistics. Although the code constraints are represented by the upper part of the graph, the channel behavior is characterized by the *channel factor nodes* at the bottom of the graph. In particular, the channel factor node relative to epoch  $k$  is connected to the variable nodes relative to the code symbols  $c_{k-\nu}^k$  (in Fig. 3,  $\nu = 2$ ) and computes the exponential metric  $\gamma_k^{\text{PLP}} (\mathbf{c}_{k-\nu}^k)$ . The application of the SP algorithm to this overall graph will allow the approximate (because of the presence of cycles in the graph) computation of the marginal code symbol *a posteriori* probabilities [29]. As it will be shown by the results presented in Section IV, the presence of short cycles containing the channel factor nodes does not entail significant performance degradation.

In the case of graph-based MAP symbol detection, the complexity can be reduced by applying the techniques proposed in [42]. In particular, a channel factor node is connected with  $\nu$  variable nodes. Although soft messages are transmitted only over the “least reliable”  $Q$  edges coming out of the channel factor node, hard messages (i.e., early decisions) are considered for the remaining  $\nu - Q$  “most reliable” edges.

### III. EXAMPLE: WIENER CHANNEL PHASE PROCESS WITH TIME-INVARIANT FREQUENCY OFFSET

In this section, we specialize the general derivation contained in the previous sections to a communication scenario characterized by an important and widely used channel model. We remark, however, that the previous derivation holds for any phase-uncertain channel, provided that the autocorrelation of the phasor process is available or can be estimated at the receiver.

We assume that the phase process  $\{\psi_k\}$  is a discrete-time Wiener process [31]. In particular, the phase evolution can be characterized by the following recursion:

$$\psi_k = \psi_{k-1} + \Delta_k \quad (29)$$

where  $\{\Delta_k\}$  are i.i.d. Gaussian increments with zero mean and variance  $\sigma_\Delta^2$ , which is descriptive of the phase noise intensity. In this case, the autocorrelation of the associated phasor process  $\{\exp(j\psi_k)\}$  becomes

$$\begin{aligned} R_\psi(n) &= \mathbb{E} \left\{ e^{j(\psi_k - \psi_{k-n})} \right\} \\ &= \mathbb{E} \left\{ e^{j(\sum_{i=0}^{n-1} \Delta_{k-i})} \right\} \\ &= \prod_{i=0}^{n-1} E\{e^{j\Delta_{k-i}}\}. \end{aligned} \quad (30)$$

It is possible to interpret  $E\{e^{j\Delta_{k-i}}\}$  as the characteristic function (evaluated in one) of the Gaussian random variable  $\Delta_{k-i}$  [31]. Since  $\{\Delta_{k-i}\}$  are identically distributed, it follows that  $E\{e^{j\Delta_{k-i}}\} = \exp(-\sigma_\Delta^2/2) \forall k, i$ . Hence

$$R_\psi(n) = \exp\left(-\frac{|n|\sigma_\Delta^2}{2}\right). \quad (31)$$

The presence of a time-invariant random uncompensated frequency offset can also be incorporated by modeling the channel phase as  $\theta_k = \psi_k + 2\pi f k T$ , where  $f$  is a random variable that is uniformly distributed in  $(-\alpha/T, \alpha/T)$ ,  $\alpha$  is the normalized frequency offset intensity, and  $T$  denotes the symbol interval. For this phase model, assuming that the phase jitter and the frequency offset are independent, it follows that

$$\begin{aligned} R_\theta(n) &= \mathbb{E} \left\{ e^{j(\theta_k - \theta_{k-n})} \right\} \\ &= \mathbb{E} \left\{ e^{j(\psi_k - \psi_{k-n})} \right\} \mathbb{E} \left\{ e^{j2\pi f n T} \right\} \\ &= \exp\left(-\frac{|n|\sigma_\Delta^2}{2}\right) \text{sinc}(2\alpha n) \end{aligned} \quad (32)$$

where  $\text{sinc}(x) \stackrel{\Delta}{=} \sin(\pi x)/\pi x$ . In the Appendix, the special case of absence of phase noise ( $\sigma_\Delta = 0$ ) and frequency offset ( $\alpha = 0$ ), i.e., the case where the channel phase reduces to a random variable uniformly distributed in  $[0, 2\pi]$ , is analyzed in more detail.

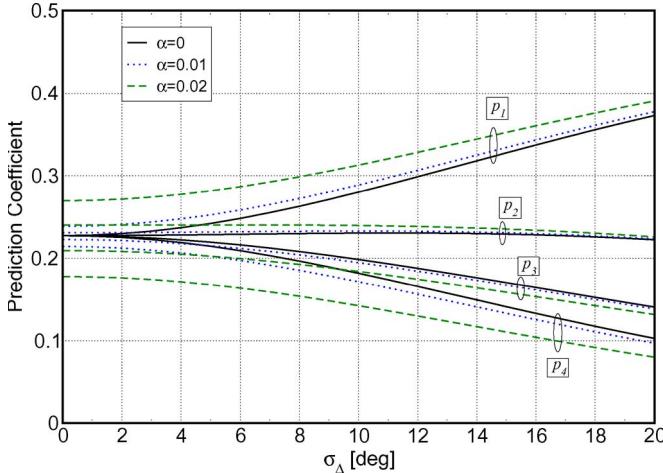


Fig. 4. Prediction coefficients as functions of the phase noise standard deviation  $\sigma_\Delta$  for an equal energy modulation, prediction order  $\nu = 4$ , and  $E_b/N_0 = 4$  dB. Various values of the frequency offset intensity  $\alpha$  are considered.

In the case of oversampling, we refer to the modified phase process  $\{\theta_k^{(i)}\}$  introduced in Section II-C for CPM signals. A detailed analysis of the effects of oversampling on the statistics of the modified phase process is rather complicated. In the case of a Wiener channel phase process  $\{\psi_k\}$ , however, a reasonable assumption consists of describing the modified phase  $\{\psi_k^{(i)}\}$  (obtained by oversampling) as a Wiener process, with an incremental variance<sup>5</sup> between  $\psi_k^{(i)}$  and  $\psi_k^{(i+1)}$  equal to  $\sigma_\Delta^2/\beta$ .

Other phase models may be considered, possibly incorporating a time-varying frequency offset, such as that caused by a Doppler shift [3]. For the sake of simplicity, we present results only for a scenario with random frequency offset and time-varying phase noise.

#### IV. NUMERICAL RESULTS

In this section, we characterize the performance of the proposed pLP detection algorithms. In particular, we first analyze the behavior of the prediction coefficients under varying channel dynamics (considering the phase process model proposed in Section III). We then evaluate the performance of the proposed receivers in terms of BER versus  $E_b/N_0$ , where  $E_b$  denotes the received energy per information bit. Several communication systems are considered in order to assess the flexibility and effectiveness of the proposed detection strategy.

Fig. 4 shows the behavior of the prediction coefficients as functions of the phase noise standard deviation  $\sigma_\Delta$  for an equal energy modulation, either phase shift keying (PSK) or CPM, a prediction order  $\nu = 4$  and  $E_b/N_0 = 4$  dB. Three values of the frequency offset intensity  $\alpha$  are considered, namely 0, 0.01, and 0.02. In the absence of phase instabilities ( $\sigma_\Delta = 0$  and  $\alpha = 0$ ), all four prediction coefficients are equal, as shown in the Appendix. For increasing phase noise or frequency offset, the prediction coefficients take on different values—the

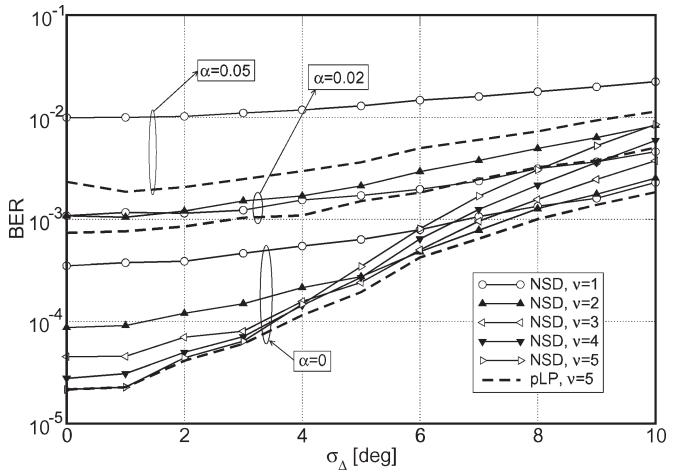


Fig. 5. BER as a function of the phase noise standard deviation  $\sigma_\Delta$  for DQPSK, symbol-by-symbol decision, and various values of the frequency offset intensity  $\alpha$ .

stronger the phase variations, the larger the difference. In particular, in the considered  $\sigma_\Delta$  range,  $p_1$  is an increasing function of  $\sigma_\Delta$ ,  $p_2$  is almost constant, whereas  $p_3$  and  $p_4$  decrease. As one can see, for no phase noise, all prediction coefficients are equal, and the proposed metric coincides with that proposed in [11]—a more detailed comparison is the subject of the Appendix. One can observe that the prediction coefficients are decreasing for an increasing value of their indexes. Schober and Gerstacker [10] consider exponentially decaying linear prediction coefficients, i.e.,  $p_i = \alpha^i$ , where  $0 \leq \alpha < 1$  is a forgetting factor. One can verify that the exponentially decaying linear prediction coefficients used in [10], which have been only heuristically justified so far, resemble the prediction coefficient profile found in our approach. Observe, however, that the exponential prediction coefficients in [10] are attractive from an implementation viewpoint.

Differentially encoded quaternary PSK (DQPSK) is considered in Fig. 5. The performance of the communication system is assessed by computer simulations in terms of BER versus the phase noise standard deviation  $\sigma_\Delta$  for various values of the frequency offset intensity  $\alpha$  and  $E_b/N_0 = 10$  dB. Symbol-by-symbol detection with decision feedback (i.e.,  $Q = 0$ ) is considered at the receiver. The performance of the proposed linear predictive receiver is compared with that of an NSD-based receiver [11]—in the latter case, the parameter  $\nu$  is related to the finite-memory parameter introduced in [11]. For  $\alpha = 0$ , the curve corresponding to the proposed pLP receiver with  $\nu = 5$  is the “envelope” of the performance curves of the NSD-based receiver relative to  $\nu = 1, 2, \dots, 5$ . For given values of  $\sigma_\Delta$  and  $\alpha$ , there exists an optimum value of  $\nu$  that minimizes the BER in the case of an NSD-based receiver—this optimal  $\nu$  reduces for increasing values of  $\sigma_\Delta$  and  $\alpha$ . On the other hand, the proposed pLP detection algorithm with  $\nu = 5$  minimizes the BER for all values of  $\sigma_\Delta$  and  $\alpha$ , provided that the prediction coefficients are adaptively updated according to the channel dynamics, i.e., the effective value of  $\sigma_\Delta$ . In the presence of frequency offset, the advantage of the proposed pLP schemes over an NSD scheme is even more pronounced, as it appears from the curves relative to  $\alpha = 0.02$  and

<sup>5</sup>Obviously, the incremental variance is the same between any two consecutive phase values inside a symbol interval and between the phases  $\psi_k^{(\beta-1)}$  and  $\psi_{k+1}^{(0)}$  across two consecutive signaling intervals.

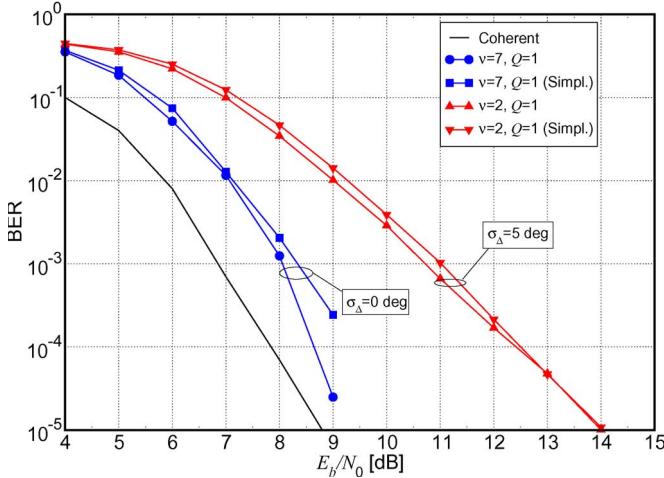


Fig. 6. BER of a TCM scheme with 16-QAM. Linear predictive receivers with various complexity levels are considered. For comparison, the performance of the equivalent coherent receiver is also shown.

$\alpha = 0.05$ . For the latter value of  $\alpha$ , the performance of the NSD schemes with  $\nu > 1$  is appreciably worse than that of the pLP schemes.

As an example of application of the proposed detection strategy to the case of nonequal energy modulations, we consider a communication system with an eight-state rate-3/4 trellis coded modulation (TCM) with 16-QAM and  $90^\circ$  rotational invariance [43]. The BER performance of the system, as shown in Fig. 6, is investigated for values of the phase jitter standard deviation  $\sigma_\Delta$  equal to  $0^\circ$  and  $5^\circ$ , respectively. For comparison, the performance with the simplified algorithm proposed in Section II-B is also evaluated. The number of states is kept fixed to 64 ( $Q = 1$ ) in all pLP receivers: This corresponds to a “strong” complexity reduction. Although for  $\sigma_\Delta = 0^\circ$ , the receiver using a prediction order  $\nu = 7$  asymptotically approaches the ideal performance of the coherent receiver (i.e., a receiver which is given perfect knowledge of the channel phase realization), for  $\sigma_\Delta = 5^\circ$ , the best performance is obtained with a small prediction order, namely  $\nu = 2$ . A possible explanation is that a large prediction order combined with heavy state reduction is likely to degrade the performance for large phase jitters. It is also worth observing that a simplified receiver based on the analysis in Section II-B shows limited performance loss with respect to the nonsimplified version. The performance of the proposed receiver could be improved, for example, by using pilot symbols. The expected improvement is not lower than that observed in communications over fading channels [44].

As an example of application of the pLP detection strategy to a transmission scheme based on CPM signaling, we consider a serially concatenated scheme characterized by an outer convolutional encoder and an inner Gaussian minimum shift keying (GMSK) modulator [45]. In particular, we refer to the Global System for Mobile Communications standard [46], where the outer code is a 16-state nonrecursive nonsystematic convolutional code with rate 1/2 and generator matrix

$$G_o(D) = [1 + D^3 + D^4 \quad 1 + D + D^3 + D^4]. \quad (33)$$

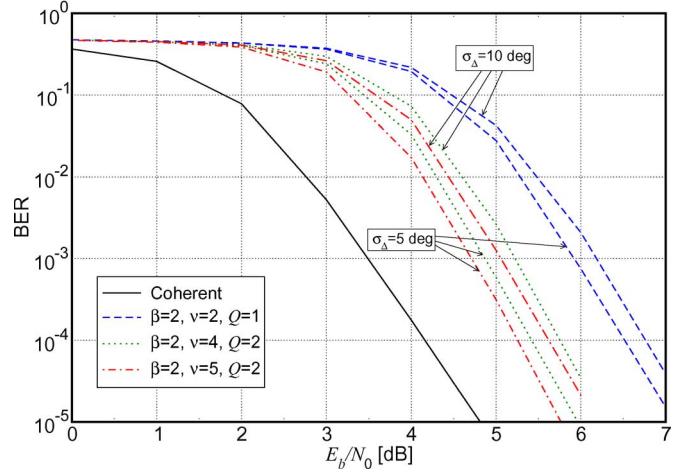


Fig. 7. BER of a serially concatenated scheme with GMSK and inner linear prediction with a sampling rate of  $\beta = 2$  samples per symbol. Various receiver complexity levels in terms of the parameters  $\nu$  and  $Q$  are considered. For comparison, the performance of the coherent system is also shown. In all cases, five decoding iterations are considered.

The outer code and the GMSK modulator are connected through a length-1024 pseudorandom bit interleaver. At the receiver side, we consider  $\beta = 2$  samples per symbol interval. The receiver is constituted by an inner detector/decoder, using a pLP-FB algorithm and relative to the CPE, and an outer decoder relative to the considered convolutional code. The numerical results are shown in Fig. 7. In all cases, five decoding iterations are considered at the receiver side. As one can see, for sufficiently large  $\nu$  and  $Q$  (for instance,  $\nu = 5$  and  $Q = 2$ ), the performance loss, with respect to an ideal coherent receiver, which perfectly knows the channel phase, is within 1.2 dB for  $\sigma_\Delta \leq 10^\circ$ .

We now describe an application of the proposed detection technique to FB algorithms and iterative detection. We consider a serially concatenated convolutional code (SCCC) consisting of an outer four-state rate-1/2 code connected through a length-1024 pseudorandom bit interleaver to an inner four-state rate-2/3 code [47]. The respective generator matrices are given by

$$G_o(D) = \begin{bmatrix} 1 & \frac{1+D^2}{1+D+D^2} \end{bmatrix} \quad G_i(D) = \begin{bmatrix} 1 & 0 & \frac{1+D^2}{1+D+D^2} \\ 0 & 1 & \frac{1+D}{1+D+D^2} \end{bmatrix}. \quad (34)$$

The output symbols are mapped to an 8-PSK constellation with natural mapping. Pilot symbols are introduced with a rate of one pilot every 16 information symbols. At the receiver side, the inner decoder uses the pLP-FB algorithm proposed in Section II-D and performs joint detection and decoding.

- 1) The behavior of the BER as a function of the SNR, for various channel dynamics, is first considered. The numerical results are shown in Fig. 8. For comparison, the performance of the corresponding ideal coherent system is also shown. In all cases, five decoding iterations between the component decoders are considered. The performance of the system for increasing phase jitter standard deviation  $\sigma_\Delta$  (from  $2^\circ$  to  $10^\circ$ ) and for various

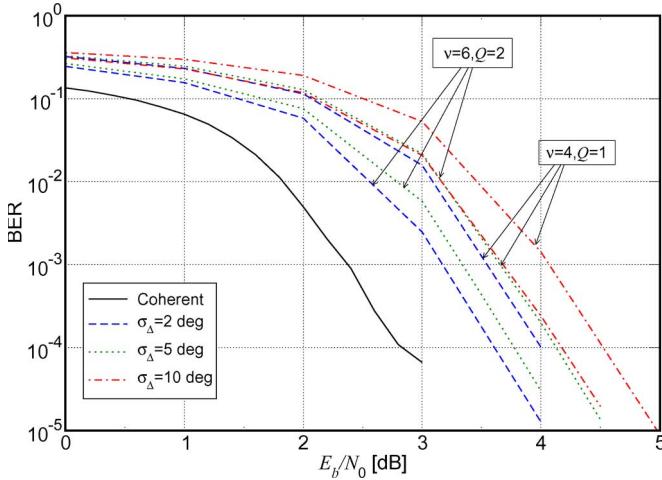


Fig. 8. BER versus SNR for an SCCC with 8-PSK and inner linear prediction at the receiver side. Various receiver complexity levels are considered. For comparison, the performance of the coherent system is also shown. In all cases, five decoding iterations are considered.

complexity levels, characterized by the couple  $(\nu, Q)$ , is evaluated. In particular, the performance loss (in terms of the SNR) with respect to the coherent system, for  $(\nu, Q) = (6, 2)$  and  $\sigma_\Delta = 2^\circ$ , is about 0.75 dB at a BER equal to  $10^{-4}$ . The loss increases to 1.4 dB for  $\sigma_\Delta = 10^\circ$ .

- 2) Next, we consider the BER performance as a function of the phase jitter standard deviation  $\sigma_\Delta$  for  $E_b/N_0 = 4$  dB. The numerical results are shown in Fig. 9. As for the results in Fig. 8, in this case as well, five decoding iterations and increasing values of the frequency offset are considered. For  $\nu = 6$  and  $Q = 3$ , the pLP-based iterative detection scheme is very robust to phase instabilities up to  $\sigma_\Delta = 10^\circ$  and  $\alpha = 0.01$ . A less complex receiver with  $\nu = 4$  and  $Q = 2$  has still acceptable performance for low values of  $\sigma_\Delta$  and  $\alpha$ . The performance of the proposed pLP-based iterative receiver degrades significantly for  $\alpha \geq 0.02$ .
- 3) Finally, we consider a direct comparison of pLP iterative detection schemes, which are proposed in this paper, with noncoherent iterative detection schemes. In particular, we refer to the noncoherent FB algorithm proposed in [47] and characterized by the following exponential metric [11]:

$$\gamma_k^{\text{NC}}(T_k) \propto P\{a_k\} \exp\left(-\frac{|c_k|^2}{N_0}\right) \frac{I_0\left(\frac{2}{N_0} |\sum_{i=0}^{\nu} r_{k-i} c_{k-i}^*|\right)}{I_0\left(\frac{2}{N_0} |\sum_{i=1}^{\nu} r_{k-i} c_{k-i}^*|\right)} \quad (35)$$

where  $I_0(x)$  is the zeroth-order modified Bessel function of the first type. As a meaningful performance comparison, we analyze the behavior of the BER as a function of the parameter  $\nu$  for a given SNR. The results are shown in Fig. 10. In particular,  $E_b/N_0 = 4$  dB, and a fixed reduced-state parameter  $Q = 2$  is considered for any value of  $\nu$ . In other words, in the case where  $\nu = 2$ , there is no complexity reduction, whereas for  $\nu \geq 3$ , the state reduction techniques in [39] are used. As one can

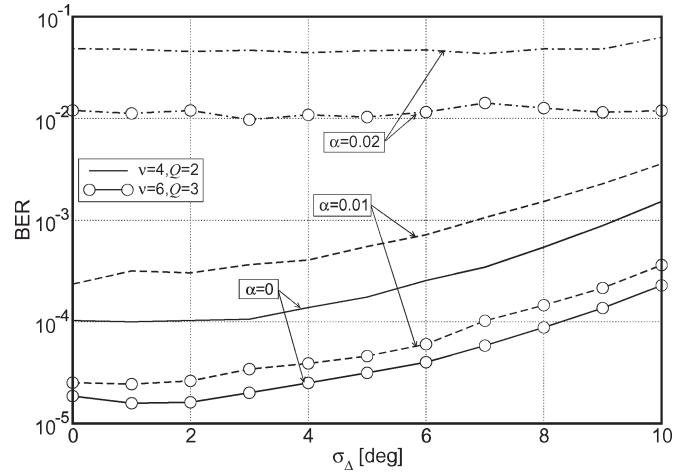


Fig. 9. BER as a function of the phase noise standard deviation  $\sigma_\Delta$  for an SCCC with 8-PSK, an inner linear predictive detector,  $E_b/N_0 = 4$  dB, various values of frequency offset intensity, and various levels of receiver complexity. In all cases, five decoding iterations are considered.

immediately see in Fig. 10, the BER of iterative detection is a nonincreasing function of the prediction order  $\nu$  for *any* phase noise standard deviation. On the contrary, the BER of noncoherent iterative detection exhibits an “optimal” value of  $\nu$ , which decreases for increasing values of the phase jitter standard deviation. The following scenarios can be distinguished.

- a) For  $\sigma_\Delta = 15^\circ$ , the performance with pLP iterative detection is better than that of noncoherent iterative detection for any value of  $\nu$ .
- b) For  $\sigma_\Delta \leq 10^\circ$ , the noncoherent iterative detection schemes outperforms the pLP iterative detection scheme for  $\nu \leq 5$ .

Considering the performance results in Fig. 5, it is possible to conclude that the use of a strong channel code (like the considered SCCC) seems to slightly favor detection strategies based on the exponential metric (35) rather than on the metric (22). We remark, however, that, by sufficiently increasing the prediction order  $\nu$ , the performance of pLP iterative detection schemes becomes identical to (and eventually outperforms) that of noncoherent iterative detection schemes. This is confirmed by the results which will be shown in Fig. 11.

We now consider the application of the proposed pLP-SP algorithm to perform graph-based joint detection and decoding of a low-density parity check (LDPC) code. In particular, we consider the transmission of a (3, 6)-regular LDPC code with codewords of length 4000 [48]. In Fig. 11, the performance of the proposed pLP-based graph-based iterative detection scheme is compared with that of the ideal coherent receiver. The considered modulation format is binary phase shift keying (BPSK), and the maximum allowed number of decoding iterations of the pLP-SP algorithm, running over a graph as in Fig. 3 with *flooding* schedule [49], is 200. A pilot symbol every 19 code symbols is added for ambiguity problems. The phase noise is modeled as a discrete-time Wiener process with incremental standard deviation over a signaling interval equal to  $\sigma_\Delta = 6^\circ, 12^\circ$ , and  $16^\circ$ , respectively. In correspondence to the

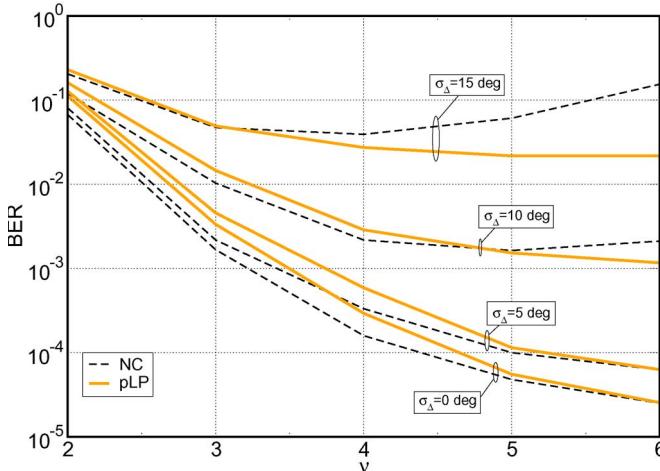


Fig. 10. BER versus parameter  $N$  in the case of iterative detection, with (solid lines) inner linear predictive and (dashed lines) noncoherent detection for an SCCC with 8-PSK. In all cases, five decoding iterations are considered, and for  $\nu \geq 3$ , the reduced-state parameter  $Q$  is fixed to 2.

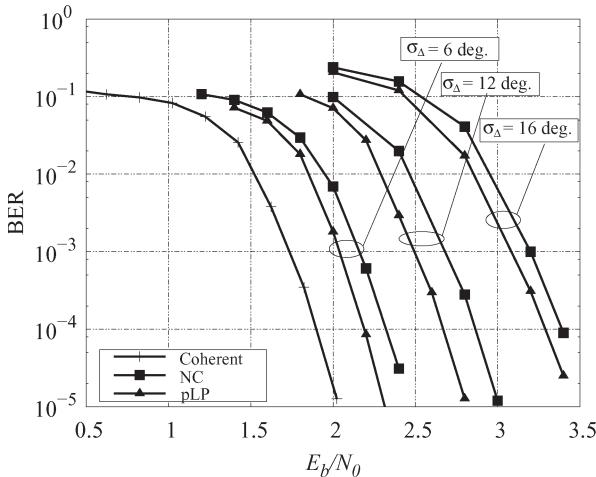


Fig. 11. BER versus SNR in the case of iterative detection based on the SP algorithm using the modified factor graph considered in Section II-E for various values of the Wiener phase noise standard deviation  $\sigma_\Delta$ . The value of  $\nu$  is optimized in each individual case.

three considered values of the phase noise standard deviation, the prediction order  $\nu$  is set to 20, 18, and 16, respectively. In all cases, a reduced-state parameter  $Q = 1$  is used: This leads to almost the same (and relatively low) complexity in all cases. As a comparison, we also consider the performance of graph-based noncoherent iterative detection, where the basic metric computed at a channel factor node has the expression (35). The parameters  $\nu$  and  $Q$  are set as in the case of graph-based pLP iterative detection. As one can see from the results in Fig. 11, the performance of the pLP receiver is slightly better than that of the noncoherent receiver.

## V. ADAPTIVITY OF pLP DETECTION ALGORITHMS

The obtained performance results with pLP detection algorithms suggest an important and interesting extension to phase-uncertain channels with time-varying dynamics. In fact, by applying standard adaptation methods, like those based on

stochastic gradient [36], it is possible to recursively “adapt” the values of the prediction coefficients (either along the survivors in a trellis diagram or in the channel factor nodes in a graph). In particular, this would make pLP detection algorithms very attractive, especially with respect to noncoherent detection algorithms, where the “sliding window” of observations used for metric computation are all weighted in the same way.

- 1) As can be seen in Figs. 5 and 10, for each specific value of the channel dynamics, there exists an optimal value of the parameter  $\nu$  for noncoherent detection algorithms. Hence, in order for noncoherent algorithms to be made adaptive, it would be necessary to dynamically vary  $\nu$ . However, this would lead to a corresponding modification of the trellis diagram used by a detection algorithm. A trellis diagram should “compress” (low values of  $\nu$ ) for fast channel variations, whereas it should “expand” (large values of  $\nu$ ) for slow channel variations. Similarly, in a factor graph, the degree of a channel factor node should reduce for low values of  $\nu$  or increase for large values of  $\nu$ . It seems complicated from an implementation standpoint that a receiver can possess such an adaptivity.
- 2) On the other hand, in Figs. 5 and 10, it can easily be seen that, by adaptively changing the prediction coefficients according to Fig. 4, increasing the prediction order improves the performance for all channel dynamics. Hence, for a fixed prediction order (and, thus, fixed receiver trellis dimension), it would be possible to consider an external block that periodically, on the basis of a short sequence of pilot symbols, estimates the current phase noise statistics and correspondingly updates the prediction coefficients. In the case of trellis-based pLP detection algorithms, the PSP principle [4] could be applied by recursively updating the prediction coefficients associated to a state from the values of the prediction coefficients at the previous state in the survivor. Adaptive versions of the pLP detection algorithms considered in this paper are currently under investigation.

We comment further on the relation between pLP detection algorithms and noncoherent algorithms. As can be seen in Fig. 4, for fast channel dynamics, the first prediction coefficient  $p_1$  dominates over the other prediction coefficients. However, the second prediction coefficient seems almost constant with respect to the phase noise standard deviation  $\sigma_\Delta$ . This is basically equivalent to considering a reduced “effective prediction window,” i.e., it is equivalent to a reduction of the observation window characterizing a noncoherent algorithm. In other words, whereas adaptive noncoherent algorithms would reduce the observation window (weighting in the same way the channel observations falling inside it), adaptive pLP algorithms would keep the same observation window but weight different channel observations, depending on the channel dynamics.

## VI. CONCLUSION

In this paper, pLP detection algorithms for phase-uncertain communications have been proposed. They are based on the application of linear predictive estimation of the channel

phasor and exploit the correlation between consecutive channel phase realizations. In particular, pLP detection algorithms have been derived for linear coded modulation and extended, by means of oversampling, to CPM signals. Both trellis-based MAP sequence detection and trellis-based and graph-based MAP symbol detection schemes have been considered. The designed pLP detection algorithms are very robust to phase noise and uncompensated frequency offset, and the design of adaptive versions is an attractive research direction. From a theoretical viewpoint, the obtained results show that linear predictive detection, although suboptimal in principle, appears to be effective even with non-Gaussian observables.

## APPENDIX TIME-INVARIANT CHANNEL PHASE WITH UNIFORM DISTRIBUTION IN $[0, 2\pi)$

The case of a time-invariant (TI) random channel phase can be derived from the Wiener phase process model considered in Section III by imposing  $\sigma_\Delta = 0$ . Assuming equal energy modulation ( $|c_k| = 1, \forall c_k$ ), the Wiener–Hopf system can be written as  $\mathbf{R}^{\text{TI}} \mathbf{p}^{\text{TI}} = \mathbf{b}^{\text{TI}}$ , with

$$\mathbf{R}^{\text{TI}} \triangleq N_0 \mathbf{I} + \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \quad (36)$$

where  $\mathbf{I}$  is the identity matrix, and  $\mathbf{b}^{\text{TI}} \triangleq [1 \cdots 1]^T$ . After a few simple manipulations, it follows that the solution of the Wiener–Hopf system consists of a sequence of identical prediction coefficients, i.e.,

$$p_i^{\text{TI}} = \frac{1}{\nu + N_0}, \quad i \in \{1, \dots, \nu\}. \quad (37)$$

Hence, the phasor estimate becomes

$$\widehat{e^{j\theta_k}} = \frac{\sum_{i=1}^{\nu} r'_{k-i}}{\left| \sum_{i=1}^{\nu} r'_{k-i} \right|} = \frac{\sum_{i=1}^{\nu} \frac{r_{k-i}}{c_{k-i}}}{\left| \sum_{i=1}^{\nu} \frac{r_{k-i}}{c_{k-i}} \right|}. \quad (38)$$

The branch metric of a pLP-VA can therefore be written, in the case of equally likely information symbols, as follows:

$$\begin{aligned} \lambda_k^{\text{PPLP-TI}}(T_k) &= -|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^{\nu} r'_{k-i}}{\left| \sum_{i=1}^{\nu} r'_{k-i} \right|} \right| \\ &= -|c_k|^2 \left| r'_k - \frac{\sum_{i=1}^{\nu} \frac{r_{k-i}}{c_{k-i}}}{\left| \sum_{i=1}^{\nu} \frac{r_{k-i}}{c_{k-i}} \right|} \right|^2. \end{aligned} \quad (39)$$

In [47], it is shown that the basic metric used in [11] for a noncoherent VA can be expressed, in the case of equally likely information symbols, as

$$\lambda_k^{\text{NSD}}(T_k) = -\Re \left\{ r_k c_k^* e^{-j\hat{\theta}_k^{(\nu)}} \right\} + \frac{|c_k|^2}{2} \quad (40)$$

where

$$\hat{\theta}_k^{(\nu)} \triangleq \arg \left( \sum_{i=1}^{\nu} r_{k-i} c_{k-i}^* \right). \quad (41)$$

After simple manipulations, it can easily be concluded that the branch metric (40) coincides with (39) in the case of equal energy modulation ( $|c_k| = 1$ ). More generally, the metric (39) coincides with the metric of the noncoherent VA proposed in [26] for QAM signals. Hence, this confirms that the proposed pLP detection algorithms generalize noncoherent detection algorithms, in the sense that the latter can be interpreted as special cases of the former.

Considering a generic modulation format (not necessarily with constant amplitude), for a large SNR, i.e.,  $N_0/|c_k|^2 \simeq 0$ , the system matrix  $\mathbf{R}^{\text{TI}}$  can be approximated as follows:

$$\begin{aligned} \mathbf{R}^{\text{TI}} &= \begin{bmatrix} \frac{N_0}{|c_{k-1}|^2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{N_0}{|c_{k-\nu}|^2} \end{bmatrix} + \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \\ &\simeq \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \end{aligned} \quad (42)$$

which implies  $p_i \simeq 1/\nu, i \in \{1, \dots, \nu\}$ . The same phasor estimate given by (38) (“exact” in the case of equal energy modulation) can therefore be approximately used for any linear modulation format, and the corresponding performance is expected to be good at a sufficiently large SNR. Hence, one can conclude that, for a large SNR, the proposed class of detection algorithms generalizes (to time-varying phase channels) the noncoherent detection strategy proposed in [26], regardless of the used modulation format.

## REFERENCES

- [1] U. Mengali and A. N. D’Andrea, *Synchronization Techniques for Digital Receivers (Applications of Communications Theory)*. New York: Plenum, 1997.
- [2] T. Pratt, C. W. Bostian, and J. E. Allnutt, *Satellite Communications*, 2nd ed. New York: Wiley, 2002.
- [3] F. Giannetti, M. Luise, and R. Reggiani, “Simple carrier frequency rate-of-change estimators,” *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1310–1314, Sep. 1999.
- [4] R. Raheli, A. Polydoros, and C. Tzou, “Per-survivor processing: A general approach to MLSE in uncertain environments,” *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 354–364, Feb.–Apr. 1995.
- [5] A. N. D’Andrea, U. Mengali, and G. M. Vitetta, “Approximate ML decoding of coded PSK with no explicit carrier phase reference,” *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1033–1039, Feb.–Apr. 1994.
- [6] A. J. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 260–269, Apr. 1967.
- [7] G. D. Forney, Jr., “The Viterbi algorithm,” *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [8] A. Anastopoulos and K. M. Chugg, “Adaptive iterative detection for phase tracking in turbo coded systems,” *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2135–2144, Dec. 2001.
- [9] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
- [10] R. Schober and W. H. Gerstacker, “Metric for noncoherent sequence estimation,” *Electron. Lett.*, vol. 35, no. 25, pp. 2178–2179, Dec. 9, 1999.

- [11] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1376–1385, Sep. 1999.
- [12] D. Divsalar and M. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 1300–1308, Mar. 1990.
- [13] D. Raphaeli, "Noncoherent coded modulation," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 172–183, Feb. 1996.
- [14] G. Colavolpe, G. Ferrari, and R. Raheli, "Noncoherent iterative (turbo) detection," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1488–1498, Sep. 2000.
- [15] G. Ferrari, G. Colavolpe, and R. Raheli, "Noncoherent iterative decoding of spectrally efficient coded modulations," *Ann. Telecommun.*, vol. 56, no. 7/8, pp. 409–421, Jul./Aug. 2001.
- [16] G. Ferrari, G. Colavolpe, and R. Raheli, "A unified framework for finite-memory detection," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1697–1706, Sep. 2005.
- [17] M. Peleg, S. Shamai (Shitz), and S. Galán, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel," *Proc. Inst. Electr. Eng. Commun.*, vol. 147, no. 2, pp. 87–95, Apr. 2000.
- [18] L. Li and A. J. Goldsmith, "Low-complexity maximum-likelihood detection of coded signals sent over finite-state Markov channels," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 524–531, Apr. 2002.
- [19] J. Lodge and M. Moher, "Maximum likelihood estimation of CPM signals transmitted over Rayleigh flat fading channels," *IEEE Trans. Commun.*, vol. 38, no. 6, pp. 787–794, Jun. 1990.
- [20] D. Makrakis, P. T. Mathiopoulos, and D. Bouras, "Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN: A combined envelope, multiple differential and coherent detection approach," *IEEE Trans. Commun.*, vol. 42, no. 1, pp. 63–75, Jan. 1994.
- [21] X. Yu and S. Pasupathy, "Innovations-based MLSE for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1534–1544, Feb.–Apr. 1995.
- [22] G. M. Vitetta and D. P. Taylor, "Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 43, no. 11, pp. 2750–2758, Nov. 1995.
- [23] F. Adachi, "Adaptive differential detection using linear prediction for M-ary DPSK," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 909–918, Aug. 1998.
- [24] L. H.-J. Lampe and R. Schober, "Low-complexity iterative demodulation for noncoherent coded transmission over Ricean-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1481–1496, Nov. 2001.
- [25] L. H.-J. Lampe and R. Schober, "Iterative decision-feedback demodulation of bit-interleaved coded MDPSK for flat Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 7, pp. 1176–1184, Jul. 2001.
- [26] G. Colavolpe and R. Raheli, "On noncoherent sequence detection of coded QAM," *IEEE Commun. Lett.*, vol. 2, no. 8, pp. 211–213, Aug. 1998.
- [27] T. Aulin and C.-E. W. Sundberg, "Continuous phase modulation—Part I: Full response signaling," *IEEE Trans. Commun.*, vol. COM-29, no. 3, pp. 196–209, Mar. 1981.
- [28] T. Aulin, N. Rydbeck, and C.-E. W. Sundberg, "Continuous phase modulation—Part II: Partial response signaling," *IEEE Trans. Commun.*, vol. COM-29, no. 3, pp. 210–225, Mar. 1981.
- [29] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [30] M. Morelli and U. Mengali, "Feedforward frequency estimation for PSK: A tutorial review," *Eur. Trans. Telecommun.*, vol. 9, no. 2, pp. 103–108, Jan.–Apr. 1998.
- [31] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1991.
- [32] C. R. Cahn, "Phase tracking and demodulation with delay," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 50–58, Jan. 1974.
- [33] P. R. Chevillat and E. Eleftheriou, "Decoding of Trellis-encoded signals in the presence of intersymbol interference and noise," in *Proc. IEEE ICC*, Philadelphia, PA, Jun. 1988, pp. 23.1.1–23.1.6.
- [34] M. V. Eyuboglu and S. U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.
- [35] A. Duel-Hallen and C. Heegard, "Delayed decision feedback estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.
- [36] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [37] B. E. Rimoldi, "A decomposition approach to CPM," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 260–270, Mar. 1988.
- [38] H. Meyr, M. Oerder, and A. Polydoros, "On sampling rate, analog pre-filtering, and sufficient statistics for digital receivers," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3208–3214, Dec. 1994.
- [39] G. Colavolpe, G. Ferrari, and R. Raheli, "Reduced-state BCJR-type algorithms," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 848–859, May 2001.
- [40] P. Thienniboon, G. Ferrari, and K. M. Chugg, "Generalized trellis-based reduced-state soft-input/soft-output algorithms," in *Proc. IEEE ICC*, New York, Apr. 2002, pp. 1667–1671.
- [41] G. Ferrari, G. Colavolpe, and R. Raheli, "On trellis-based truncated-memory detection," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1462–1476, Sep. 2005.
- [42] G. Colavolpe, "On LDPC codes over channels with memory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1757–1766, Jul. 2006.
- [43] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*. New York: Macmillan, 1991.
- [44] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channel," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.
- [45] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
- [46] ETSI, *ETSI EN300 909 v.8.3*, Jul. 2000.
- [47] G. Ferrari, A. Anastasopoulos, G. Colavolpe, and R. Raheli, "Adaptive iterative detection for the phase uncertain channel: Limited-tree-search versus truncated-memory detection," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 433–442, Mar. 2004.
- [48] D. J. C. MacKay, *Regular LDPC Online Database*. [Online]. Available: <http://www.inference.phy.cam.ac.uk/mackay/>
- [49] F. Kschischang and B. Frey, "Iterative decoding of compound codes by probability propagation in graphical models," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 219–230, Feb. 1998.



**Gianluigi Ferrari** (S'97–M'03) was born in Parma, Italy, in November 1974. He received the Laurea (five-year program, *summa cum laude*) and Ph.D. degrees in electrical engineering from the University of Parma, in October 1998 and January 2002, respectively.

From July 2000 to December 2001, he was a Visiting Scholar with the Communication Sciences Institute, University of Southern California, Los Angeles. Since 2002, he has been a Research Professor with the Department of Information Engineering, University of Parma. Between 2002 and 2004, he visited several times, as a Research Associate, the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA. He has published more than 70 papers in leading international journals and conferences. He is a coauthor of the books *Detection Algorithms for Wireless Communications, With Applications to Wired and Storage Systems* (Wiley, 2004), *Teoria della probabilità e variabili aleatorie con applicazioni* (McGraw-Hill, 2005), and *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective* (Wiley, 2006). His research interests include digital communication systems design, adaptive signal processing for wireless and optical communications, information theory, and *ad hoc* wireless networking.

Dr. Ferrari coreceived a Best Student Paper Award at the 2006 IEEE International Workshop on Wireless *Ad hoc* and Sensor Networks (IWAN'06). He acts as a frequent reviewer for many international journals and conferences and is also a Technical Program Member for several international conferences. Since 2007, he has been on the Editorial Board of the *Open Electrical and Electronic Engineering (OEEE) Journal*, Bentham Publishers.



**Giulio Colavolpe** (S'96–M'00) was born in Cosenza, Italy, in 1969. He received the Dr. Ing. degree in telecommunications engineering (*cum laude*) from the University of Pisa, Pisa, Italy, in 1994 and the Ph.D. degree in information technology from the University of Parma, Parma, Italy, in 1998.

Since 1997, he has been with the University of Parma, where he is currently an Associate Professor of telecommunications with the Department of Information Engineering. In 2000, he was a Visiting Scientist at the Institut Érecom, Valbonne, France.

He is also the Principal Investigator of several research projects funded by the European Space Agency (ESA-ESTEC) and important telecommunications companies. His main research interests include digital transmission theory, channel coding, and signal processing. His research activity has led to several scientific publications in leading international journals and conference proceedings and a few industrial patents. He is also a coauthor of the book *Detection Algorithms for Wireless Communications* (Wiley).

Dr. Colavolpe received the Best Paper Award for the paper “VLSI design of a fully parallel decoder for LDPC-like codes” at the 13th International Conference on Software, Telecommunications, and Computer Networks (SoftCOM 2005) in 2005.



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

From 1986 to 1988, he was with Siemens Telecomunicazioni, Cassina de’ Pecchi, Milan, Italy. From 1988 to 1991, he was a Research Professor with 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 communications engineering within the Department of Information Engineering. His scientific interests are in the general area of statistical communication theory, with application to wireless, wired, and storage systems and special attention to data detection in uncertain environments, iterative information processing, and adaptive algorithms for communications. His research has led to numerous scientific publications in leading international journals and conference proceedings, as well as a few industrial patents. He is a coauthor of the book *Detection Algorithms for Wireless Communications, With Applications to Wired and Storage Systems* (Wiley, 2004). In 1990, he conceived (with A. Polydoros) the principle of *Per-Survivor Processing*. Since 2003, he has been on the Editorial Board of the EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS as an Editor for Communication Theory.

Dr. Raheli has been on the Editorial Board of the *European Transactions on Telecommunications* as an Editor for Communication Theory since 2003. Previously, he served on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS as an Editor for Detection, Equalization, and Coding from 1999 to 2003. He also served as a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, *Special Issue on Differential and Noncoherent Wireless Communications*, which was published in September 2005.