Detection by Multiple Trellises

M. Franceschini, *Member, IEEE*, G. Ferrari, *Member, IEEE*, and R. Raheli, *Member, IEEE*

***Abstract*—In this paper, we present a novel pragmatic approach, referred to as *detection by multiple trellises*, to perform trellis-based detection over realistic channels. More precisely, we consider channels with unknown parameters and apply the concept of detection by multiple trellises to forward-backward (FB) algorithms. The key idea of our approach consists, *first*, of properly quantizing the channel parameters and, *then*, considering replication of *coherent* FB algorithms operating on parallel trellises, one per hypothetical quantized value. In order to make the receiver robust against a possibly time-varying channel parameters, the proposed soft-output algorithms perform a proper "manipulation" of the forward and backward metrics computed by the parallel FB algorithms at regularly spaced trellis steps. We consider two significant examples of application: detection over (i) phase-uncertain channels and (ii) fading channels. The performance of the proposed algorithms is investigated considering differentially encoded (DE) quaternary phase shift keying (QPSK) and iterative detection schemes based on low-density parity-check (LDPC) codes. Besides having a low complexity, the proposed soft-output algorithms turn out to be robust, flexible, *blind*, in the sense that no knowledge of the channel parameter statistics is required, and *highly parallelizable*, as it is desirable in high-throughput future wireless communication systems.**

***Index Terms*—Forward backwards algorithm, non-coherent detection, soft-input soft-output detection, LDPC codes.**

I. Introduction

DETECTION over channels which depend on unknown parameters, i.e., detection with unknown channel state information (CSI), has long been an active research field in the literature. Several signal processing techniques have been developed in the last decades to overcome possible impairments of the communication channels. Since the introduction of turbo-codes (TC) more than a decade ago [1], a great effort has been devoted to develop soft-input soft-output (SISO) detection algorithms suitable for iterative processing [2]. While SISO algorithms were first derived for the additive white Gaussian noise (AWGN) channel, they have been extended to more realistic channels, such as those of interest in wireless communications. In particular, these channels are often characterized by time-varying parameters, either stochastic or deterministic but unknown. In addition, the statistics of the stochastic parameters may not be available at the receiver. An example of such channels is the phase-uncertain channel [3], where the transmitted signal undergoes an unknown phase rotation and is affected by AWGN. Another relevant example is a fading channel [4], which arises because of unresolvable multipath in radio communications.

Two main approaches to perform detection over channels with parametric uncertainty can be devised:

- *separate* detection and parameter estimation [3];
- *joint* detection and parameter estimation. In the case of phase-uncertain communications, parameter estimation may be embedded in the detection process, explicitly [5] or implicitly [6]–[9].

In [10]–[14], linear predictive receivers for fading channels are proposed, considering the Clarke model for fading channels [15], [16]: these receivers exploit the correlation characteristics of the fading process to predict its evolution. Another general approach consists of describing the evolution of the fading process through a suitable Markov chain [17]–[19], and then taking this model into account in the receiver design [20]–[24]. A major issue, in the design of SISO algorithms for realistic channels, is to obtain good performance together with high robustness against channel variations and low computational complexity.

In this paper, we present a class of low-complexity SISO algorithms, derived from the standard forward-backward (FB) algorithm[1] [25]. These algorithms stem from an optimal approach to detection for channels affected by block-constant time-varying unknown parameters. In particular, after a proper quantization of the channel parameters, the proposed algorithms consist of running a number of coherent standard FB algorithms in parallel, which exchange information only at a relatively small number of fixed trellis epochs. As relevant case studies, we focus on differentially encoded (DE) quaternary phase shift keying (QPSK) over (i) phase-uncertain and (ii) flat fading channels. We focus our attention on the performance of the proposed algorithms in the low signal-to-noise ratio (SNR) region, which is of interest in modern communication systems. As a consequence, our analysis focuses, besides on a simple DE-QPSK scheme, also on schemes based on the concatenation of a low-density parity-check (LDPC) code [26] with DE-QPSK. In both phase-uncertain and fading channels, the obtained results show that the proposed algorithms have good robustness and low sensitivity to the statistics of the channel parameters. In the fading scenario, we compare the performance of the proposed approach with that obtained by modeling the fading process by a first-order

Paper approved by H. Leib, the Editor for Communication and Information Theory of the IEEE Communications Society. Manuscript received April 28, 2007; revised August 14, 2007.

M. Franceschini was with the Department of Information Engineering, University of Parma, Italy. He is now with the IBM T. J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598 (e-mail: franceschini@us.ibm.com).

G. Ferrari and R. Raheli are with the Department of Information Engineering, University of Parma, viale G. P. Usberti 181A, I-43100 Parma, Italy (e-mail: {gianluigi.ferrari, raheli}@unipr.it).

This paper was presented in part at the *IEEE Intern. Symp. Inform. Theory* (ISIT'05), Adelaide, Australia, Sept. 2005, and the *Intern. Symp. of Wireless Commun. Syst. 2005* (ISWCS 2005), Siena, Italy, Sept. 2005.

Digital Object Identifier 10.1109/TCOMM.2009.03.070185

[1] Also widely known as BCJR algorithm from the initials of the original proposers [25].

Markov-chain [17]–[19]. Finally, we show that by applying the principle of detection by multiple trellises, it is possible to obtain low-complexity soft-output algorithms with negligible impact on the system performance. In addition, the proposed algorithms do not rely on any statistical information on the channel parameters (either phase or fading), i.e., they are *blind*. We note that a multiple-trellis approach was previously proposed to cope with the problem of joint detection and modulation classification in [27], where the unknown parameter (modulation format) was inherently static.

This paper is organized as follows. In Section II, we introduce the idea of detection by multiple trellises. In Section III, we apply the multi-trellis SISO algorithms to DE phase-uncertain communications and analyze, by computer simulations, the performance of the proposed algorithms. In Section IV, we extend the derivation proposed in the previous section to flat fading communications and compare the results with a Markov chain-based approach. A simple complexity analysis is presented in Section V, comparing the proposed algorithm with known solutions. Section VI concludes the paper. In the Appendix, a formulation of an FB algorithm for transmission through a Markov chain channel is given.

II. Detection by Multiple Trellises: The Idea

In order to set the problem under study and present the mathematical notation, we begin by reviewing a modified version of the FB algorithm suitable for generic finite-memory channels affected by time-invariant stochastic parameters. Afterwards, we will describe the extension to time-varying parameters and propose two different multi-trellis SISO algorithms.

A. Time-Invariant Parameters

Let us assume that the channel output is observed for a period of $K+1$ symbol intervals. The channel can be completely described by the following joint probability density function (PDF)

$$p(\boldsymbol{r}_0^K, \xi | \boldsymbol{a}_0^K) \tag{1}$$

where $\boldsymbol{r}_0^K$ is the vector of the observables $(r_0, \ldots, r_K)$, $\xi \in \mathcal{D}_\xi$ is a stochastic channel parameter independent of the transmitted data, $\mathcal{D}_\xi$ is the domain of the channel parameter, and $\boldsymbol{a}_0^K$ is the vector of information symbols a_k transmitted through this channel. Note that (1) can take into account possible coding of the information symbol sequence $\{a_k\}$ into a code sequence $\{c_k\}$. We remark that the parameter ξ could be either a scalar parameter or a vector parameter, i.e., ξ could represent a whole set of parameters.

The *a posteriori* probability (APP) of an information symbol a_k can be expressed as follows:

$$\begin{aligned} P\{a_k | \boldsymbol{r}_0^K\} &\propto p(\boldsymbol{r}_0^K | a_k) P\{a_k\} \\ &= P\{a_k\} \int_{\mathcal{D}_\xi} p(\boldsymbol{r}_0^K | a_k, \xi) p(\xi)\, \mathrm{d}\xi \end{aligned} \tag{2}$$

where the notation "$\propto$" indicates that the first member is proportional to the second through a constant independent of the transmitted information symbol a_k. If, conditionally on the parameter realization ξ, the channel has finite memory [9], the conditional PDF $p(\boldsymbol{r}_0^K | a_k, \xi)$ can be computed via a standard FB algorithm [2], [25]. This is possible whenever the transmission system can be modeled as a finite state machine (FSM) whose input and output are, respectively, the information symbol a_k and a random variable (RV) whose statistics depend only on the FSM state and the input symbol.

A simple approximation for the computation of the integral in (2) consists of performing the following finite sum:

$$P\{a_k | \boldsymbol{r}_0^K\} \mathrel{\tilde{\propto}} P\{a_k\} \sum_{i=1}^{L} p\left(\boldsymbol{r}_0^K | a_k, \xi^{(i)}\right) p(\xi^{(i)}) \tag{3}$$

where $\{\xi^{(1)}, \ldots, \xi^{(L)}\}$ is a set of quantized values for the channel parameter whose positions and number L are chosen to obtain the desired accuracy in the numerical integration in (2). This corresponds to running L standard FB algorithms (each one associated with a value $\xi^{(i)}$, $i = 1, \ldots, L$) in parallel and computing a weighted average of their outputs to obtain a quantity approximately proportional to the APP.[2]

In the following, we denote the *forward state metrics* computed during the forward recursion of an FB algorithm as $\{\alpha^{(i)}(s_k)\}$, where the superscript i refers to the FB algorithm associated with the quantized parameter value $\xi^{(i)}$ and s_k denotes the state of the FSM in the corresponding trellis diagram. In particular, we assume that $s_k \in \{0, \ldots, \Xi - 1\}$, where Ξ is the number of states characterizing each trellis. Similarly, we denote the *backward state metrics* associated with the i-th trellis diagram as $\{\beta^{(i)}(s_k)\}$.

Several practical scenarios can be cast within the model described by (1), (2) and (3). In particular, as useful examples, we will consider phase-uncertain and flat fading channels.

1) Phase-Uncertain Channel: In a communication scenario where the channel introduces a time-invariant phase rotation, the stochastic channel parameter ξ can be equivalently modeled as a phase rotation θ of the transmitted symbol sequence. The discrete-time equivalent observation can be expressed as

$$r_k = c_k e^{j\theta} + n_k \tag{4}$$

where r_k is the received observable, c_k is the (possibly coded) transmitted symbol, and n_k is a (noise) sample of a sequence of independent and identically distributed (i.i.d.) zero mean Gaussian RVs.

2) Flat Fading Channel: The generic observation model given by (1) applies directly to a flat fading channel, provided that ξ has the proper statistical distribution. In particular, in a scenario with unresolvable multipath, ξ corresponds to a fading coefficient f and the channel input-output relation can be expressed as follows:

$$r_k = f\, c_k + n_k \tag{5}$$

where, in the case of Rayleigh fading, f has a complex circularly-symmetric Gaussian distribution with zero mean.

B. Time-Varying Parameter

The idea of detection by multiple trellises stems from an extension of the previous *static-parameter* approach to a scenario with *time-varying* channel parameters.

[2] We implicitly assume that the reader is familiar with the FB algorithm. More information can be found in [2], [25].

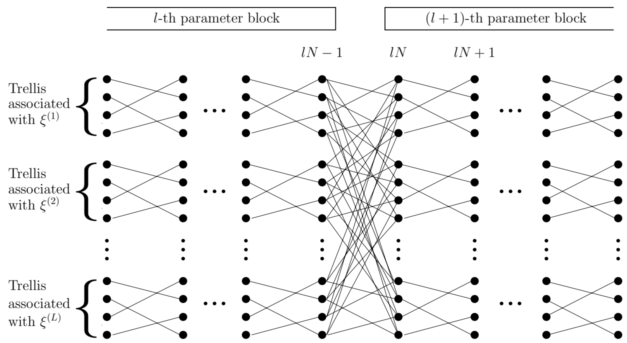

Fig. 1. Time-varying trellis for detection on block-constant discrete parameter channel.

In order to obtain insights on the impact of the presence of a time-varying parameter, let us consider a useful case study where the channel parameter process $\{\xi_k\}$ is discrete and block constant. Let us assume that ξ_k is uniformly distributed over the set $\{\xi^{(1)}, \ldots, \xi^{(L)}\}$ and constant over blocks of length $N < K$. In other words,

$$\xi_{lN+i} = \xi_{lN+j} \qquad \forall i, j \in \{0, \ldots, N-1\}$$

and the realizations $\{\xi_k\}$ are i.i.d. from block to block, i.e.,

$$p(\xi_{lN}, \xi_{nN}) = p(\xi_{lN})p(\xi_{nN}) = \frac{1}{L^2} \qquad \forall l \neq n\,.$$

As a consequence, the process $\{\xi_k\}$ is a time-varying Markov chain, characterized by an $L \times L$ transition matrix $P_k = (p_{ij}^{(k)})$ at the k-th epoch such that

$$p_{ij}^{(k)} = \begin{cases} \delta_{ij} & \text{if} \quad k \neq N-1 \mod N \\ \dfrac{1}{L} & \text{if} \quad k = N-1 \mod N \end{cases}$$

where δ_{ij} denotes the Kronecker delta. We further assume that the information sequence $\{a_k\}$ is encoded into a code symbol sequence $\{c_k\}$ by means of an FSM. Considering that the observed sequence of length K comprises more than one length-N block with constant channel parameter, the application of a maximum a posteriori (MAP) strategy to this scenario leads to a time varying trellis. In the Appendix, a general formulation of a MAP algorithm for a finite-memory channel characterized by a generic Markov-chain parameter is presented. In Fig. 1, a representative time-varying trellis for this illustrative block-constant discrete parameter channel is shown. Within a block, i.e., for $N-1$ consecutive time epochs, the trellis structure consists of L "coherent" trellises, each assuming knowledge of ξ, one for each quantized value of ξ. In the sections of the various trellis diagrams connecting the states at the end of a block with the states at the beginning of the next block, each state in each coherent trellis is connected with the corresponding state in all the other coherent trellises. In other words, each coherent trellis is connected with any other trellis by the non-zero probability of variation of the parameter value.

Applying the general formulation in the Appendix, the forward and backward metrics $\alpha_k(s_k, \xi_k)$ and $\beta_k(s_k, \xi_k)$ are functions of the "extended" state $\sigma_k = (s_k, \xi_k)$. They can be computed recursively by running L separate coherent FB algorithms, one for each parameter value. Every N time steps, in general, $\alpha_k(s_{k+1}, \xi_{k+1})$ and $\beta_k(s_k, \xi_k)$ depend on all forward and backward metrics in all coherent trellises, respectively, i.e., a "mix" of the forward and backward metrics in the coherent FB algorithms is performed. The above considerations can be equivalently drawn by following the guidelines in [21], where a Markov-chain model for the channel phase is assumed.

At this point, the idea underlying detection by multiple trellises can be outlined. As for a constant channel parameter ξ, several coherent FB algorithms are run independently, characterized by forward and backward metrics $\alpha_k^{(i)}(s_k) = \alpha_k(s_k, \xi^{(i)})$ and $\beta_k^{(i)}(s_k) = \beta_k(s_k, \xi^{(i)})$, respectively. The difference with respect to the time-invariant channel parameter case is that every N time steps, the forward (backward) metrics in the different trellises are properly "mixed" to account for the possible variation of the channel parameter. In the following, we will refer to N as "inter-mix interval."

The idea of considering parallel trellises which occasionally "talk" to each other is appealing, since it is likely to allow both low-complexity and parallel processing. In Section V, we will investigate the complexity of the proposed algorithms and compare them with other existing solutions. In this sense, performing detection by multiple trellises can be equivalently interpreted as an instance of the *divide et impera* approach to tackle complicated problems with limited complexity.

The "mix strategy," in general, should be tailored for the specific communication scenario at hand. Nevertheless, some general considerations can be drawn:

- If ξ is time-invariant, the quantity $p(\boldsymbol{r}_0^K | a_k, \xi^{(i)})$, computed via a coherent FB algorithm, is expected to be maximum in correspondence of the value $\xi^{(i)}$ closest to the true[3] channel parameter ξ. In fact, numerical analysis in several scenarios showed that the forward and backward state metrics $\{\alpha_k^{(i)}(s_k)\}$ and $\{\beta_k^{(i)}(s_k)\}$ exhibit an exponential decay[4] as a function of the epoch k. This is due to the fact that, if $\boldsymbol{\alpha}_k^{(i)}$ is the vector of the forward metrics at epoch k in the i-th trellis, the forward recursion can be equivalently expressed as
$$\boldsymbol{\alpha}_k^{(i)} = \Gamma_{k-1}^{(i)} \boldsymbol{\alpha}_{k-1}^{(i)} \qquad (6)$$
where $\Gamma_k^{(i)}$ is a matrix whose elements are the pdfs of the observable r_k conditioned to every possible transitions in the i-th coherent trellis. In particular, as expected, the decay exponent is greater in the FB algorithm associated with the phase value $\xi^{(i)}$ which is closest to the *true* channel parameter ξ, leading to state metrics $\{\alpha_k^{(i)}(s_k)\}$ and $\{\beta_k^{(i)}(s_k)\}$ relatively much larger than those computed by the j-th FB algorithm with $j \neq i$.
- If ξ is time-varying, we expect that $\{\alpha_k^{(i)}(s_k)\}$ and $\{\beta_k^{(i)}(s_k)\}$ will try to *adapt* to the parameter changes.

[3]Depending on the symmetry structure of the modulation code, i.e., the law encoding the information symbols a_k into the transmitted symbols c_k, there can be a *set* of ξ values which are optimal, in the sense that they are undistinguishable at the receiver. This may occur, for example, in differential M-PSK transmitted over a phase uncertain channel, where phase rotations of the observed sequences by multiples of $2\pi/M$ cannot be distinguished [8], [21].

[4]In the probability domain.

This adaptiveness is limited by the fact that state metrics exhibit a "low-pass filter" behavior, i.e., state metrics have *memory* and can change only slowly. This is due to the recursive structure of the metric computation algorithm (6). In other words, the FB metric computation process can be equivalently described as a recursive time-varying vector filtering.

- While in standard applications an FB algorithm is insensitive to a possible multiplication of all forward or backward state metrics by a constant, in the algorithm underlying (3), the relative weights of different trellises are important. Accordingly, the multi-trellis SISO algorithm turns out to be insensitive to a *normalization* of the metrics *only if* this normalization is carried out, at a given epoch, over all forward or backward state metrics of all parallel FB algorithms.

In the following, two possible "mix" strategies are proposed. These strategies will be analyzed in Section III and IV.

1) Multi-Trellis SISO Algorithm 1: At each length-N interval, i.e., at epochs $k = lN$, $l \in \mathbb{N}$, one could manipulate the forward metrics $\{\alpha_k^{(i)}(s_k)\}$ (and, similarly, the backward metrics $\{\beta_k^{(i)}(s_k)\}$) according to the following rule:

$$\alpha_k^{(i)}(s_k) \longleftarrow \sum_{j=1}^{L} \alpha_k^{(j)}(s_k) \quad i = 1, \ldots, L \quad \forall s_k \tag{7}$$

where the notation "$\longleftarrow$" represents the assignment of a new value. This corresponds to averaging, for every given state s_k, the metrics relative to all quantized phase values: in other words, the metrics associated with a given state in the various trellises are averaged. We will refer to this algorithm as Algorithm 1. This is the exact APP computation algorithm for the channel with block-constant parameter described at the beginning of Section II-B, if the observables are independent (conditionally on the parameter and the data sequence).

2) Multi-Trellis SISO Algorithm 2: Assume that the channel is slowly time-varying, i.e., we assume that ξ can exhibit small changes in adjacent epochs. If we allow a suitable manipulation of $\{\alpha_k^{(i)}(s_k)\}$ and $\{\beta_k^{(i)}(s_k)\}$ only at epoch $k = lN$, with $l \in \mathbb{N}$, the possible transitions of the parameter from one quantization interval to another, occurring amid the block, should be taken into account. Heuristically, we have discovered that the impact of slow parameter changes within the block can be limited by performing a normalization of the forward state metrics $\{\alpha_k^{(i)}(s_k)\}$ (and, similarly, of the backward state metrics $\{\beta_k^{(i)}(s_k)\}$) as follows:

$$\alpha_k^{(i)}(s_k) \longleftarrow \frac{\alpha_k^{(i)}(s_k)}{\sum_{s'_k} \alpha_k^{(i)}(s'_k)} \quad i = 1, \ldots, L \quad \forall s_k\,. \tag{8}$$

where s'_k is a dummy state in the summation, running over all Ξ states of a coherent trellis. This corresponds to a normalization of the state metrics within each FB algorithm, i.e., trellis by trellis, as opposed to a normalization amongst all trellises (as considered in Algorithm 1). We will refer to this algorithm as Algorithm 2.

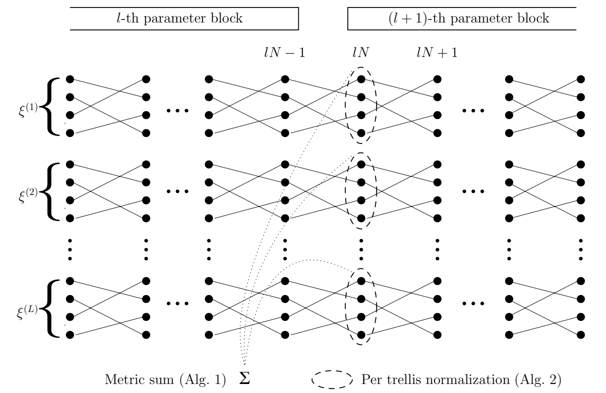

Fig. 2. Pictorial exemplification of the metric mixes in the two proposed algorithms.

3) Metric Mix in the Algorithms: a Comparison: The manipulations corresponding to (7) and (8) can be interpreted as a combining or *mixing* of the metrics $\{\alpha_k^{(i)}(s_k)\}$ (similarly for the metrics $\{\beta_k^{(i)}(s_k)\}$). Fig. 2 gives a pictorial description of the proposed algorithmic family, highlighting the *metric mix* for both Algorithms 1 and 2. Each depicted trellis diagram is associated with a coherent FB algorithm which assumes a given channel parameter $\xi^{(i)}$, $i = 1, \ldots, L$. The metric mix for Algorithm 1 is shown to "manipulate" the metrics of all trellises summing all metrics on a per-state basis, whereas the metric mix for Algorithm 2 "manipulates" each trellis independently of the other trellises, performing a per-trellis normalization. The mix position lN, i.e., the beginning of the block, refers to the forward metric computation. The backward metric computation mix is performed at epochs $lN - 1$.

In both Algorithms 1 and 2, the value of L, i.e., the number of quantized values of the channel parameter, must be chosen considering its impact on both performance and complexity. In particular, by increasing L one can improve the performance of the proposed detection algorithms, even though for sufficiently large value of L the performance improvement becomes negligible. On the other hand, as will be shown in Section V, the complexity of the detection algorithms increases linearly with L.

III. Detection by Multiple Trellises for Phase-Uncertain Channels

In this section, the phase-uncertain channel is considered. First, the algorithms introduced in Section II are specialized to this type of channel. Then, these algorithms are analyzed and numerical results are given to characterize their performance.

A. SISO Detection Algorithms

In Section II-A1, the model for a channel introducing a time-invariant phase rotation θ is given. In this case, the APP of an information symbol a_k is given by (2). Assuming that θ is uniformly distributed, i.e., $p_\theta(\vartheta) = 1/2\pi$ for $\vartheta \in [0, 2\pi)$ (and 0 otherwise), expression (3) specializes to the following:

$$P\{a_k|\boldsymbol{r}_0^K\} \overset{\sim}{\propto} P\{a_k\} \sum_{i=1}^{L} p(\boldsymbol{r}_0^K|a_k, \vartheta^{(i)}) \tag{9}$$

where $\{\vartheta^{(1)}, \ldots, \vartheta^{(L)}\}$ is a set of L *properly chosen* phase values [28]. This detection approach for channels with a block-constant random phase was used in [8].

If we assume a slowly varying channel phase (i.e., the bandwidth of the channel parameter process is small compared with the receiver filter bandwidth), the discrete-time observable can be modeled as in (4) by incorporating a time-varying phase process $\{\theta_k\}$:

$$r_k = c_k e^{j\theta_k} + n_k \ . \tag{10}$$

where $|c_k| = 1$ (DE-QPSK is considered) and n_k is a discrete-time complex AWGN process with $\mathrm{Var}\{n_k\} = (RE_\mathrm{b}/N_0)^{-1}$, in which R is the system spectral efficiency in bits per channel use. By suitably modeling the stochastic process $\{\theta_k\}$, one could try to develop an *exact* APP algorithm. Since we do not want to rely on exact channel parameter statistics, we resort to the multi-trellis SISO algorithms described in Section II-B.

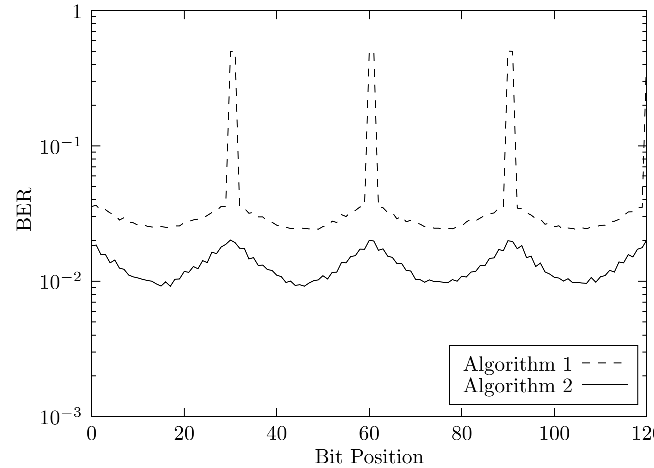

Fig. 3. BER performance at each codeword position for Algorithms 1 and 2 and DE-QPSK. $N = 15$ and $E_b/N_0 = 6.0$ dB.

B. Numerical Results

In this section, we assume that transmission over an AWGN channel is affected by a Wiener phase noise process $\{\theta_k\}$ described by the following recursive relation:

$$\theta_k = \theta_{k-1} + w_k \mod 2\pi \tag{11}$$

where $\{w_k\}$ is a sequence of i.i.d zero mean Gaussian variables. The standard deviation of w_k, denoted as σ_θ, is representative of the phase noise intensity. As mentioned in Section I, the chosen modulation format is DE-QPSK.

Both Algorithms 1 and 2 introduced in Section III are considered. For Algorithm 1, the number of quantized phase values is $L = 32$, i.e., 8 values per phase interval between adjacent QPSK symbols: in other words $\vartheta^{(i)} = 2\pi i/32$, $i = 0, \ldots, 31$. For Algorithm 2, $L = 8$ and $\vartheta^{(i)} = 2\pi i/32$, $i = 0, \ldots, 7$, accounting for a phase interval of only $\pi/2$. In fact, it is well known that for a time-invariant channel phase, the symmetry of DE-QPSK enables to perform detection accounting only for a phase interval $(0, \pi/2)$ [8], [21]. The particular structure of Algorithm 2, i.e., the fact that the normalization is carried out on a "per-coherent trellis" basis, allows to perform this complexity reduction at the cost of a limited penalty also in the presence of a time-varying channel phase.

In Fig. 3, Algorithms 1 and 2 are investigated for detection of DE-QPSK without an outer code. The information symbols are transmitted in blocks. The bit error rate (BER) performance relative to the first 120 bits in the transmitted blocks is shown, as a function of the bit position within the block. The bit SNR E_b/N_0 is equal to 6 dB. The inter-mix interval is $N = 15$ and the phase noise parameter $\sigma_\theta = 5°$. One can note that the curves exhibit periodicity $2N$, since each QPSK symbol encodes 2 bits. In particular, for both algorithms the BER is lowest at the epoch in the middle between two consecutive metric mixes. Algorithm 1, moreover, exhibits a floor, since bits corresponding to the metric mix epochs, i.e., bits at positions $l2N$ and $l2N + 1$, are randomly decided. This floor has little impact in a concatenated coded scheme, as considered in the following paragraphs, since bits at positions $l2N$ and $l2N + 1$ are characterized by APP equal to 0.5, i.e., they behave as punctured bits.

We now assume that the information sequence is encoded by an outer regular (3,6) LDPC code [26]. The codeword length is set to 6000 bits. The decoder uses a standard LDPC decoder as an external SISO module which exchanges extrinsic information with the DE-QPSK inner detector, where the proposed SISO algorithms are used instead of a coherent SISO algorithm for DE-QPSK. This can obviously be interpreted as a serially concatenated coding scheme. More details on the considered concatenated system structure can be found in [28]. The maximum number of iterations is set to 100.

In Fig. 4, the performance of the described schemes is shown in terms of BER versus SNR. The performance for transmission over an AWGN channel without phase noise, considering an ideal coherent FB algorithm as inner detector, is shown as a reference. The remaining curves show the performance obtained with the proposed algorithms. In particular, the curves marked as "Alg1" and "Alg2" correspond to the performance of the schemes with Algorithms 1 and 2, respectively. For each algorithm, several values of the phase noise standard deviation σ_θ (given in degrees in the figure legend) are considered. In each case, the inter-mix interval N is heuristically optimized. The results in Fig. 4 show that, even in the presence of a significant phase noise (for instance, $\sigma_\theta = 10°$), it is possible to "blindly" process the metrics of the trellises while still achieving an SNR loss as limited as 1 dB. Heuristically, the optimum value of N turns out to be inversely proportional to σ_θ. The results in Fig. 4 show that Algorithm 2 entails better performance than Algorithm 1. In particular, for very strong phase noise, i.e., $\sigma_\theta = 10°$, Algorithm 1 suffers an SNR penalty larger than 1 dB with respect to Algorithm 2. This is due to the fact that Algorithm 1 completely erases the phase information every N time steps, whereas Algorithm 2 performs only a "trellis balancing" as described in Section II-B2.

In Fig. 5, a direct comparison between the performance (in terms of BER as a function of the SNR) with Algorithm 1 and Algorithm 2, for a fixed value of the inter-mix distance $N = 15$, and several values of σ_θ, is shown. The value $N = 15$ optimized the system performance at $\sigma_\theta = 5°$, as shown in Fig. 4. The remaining system and simulation parameters

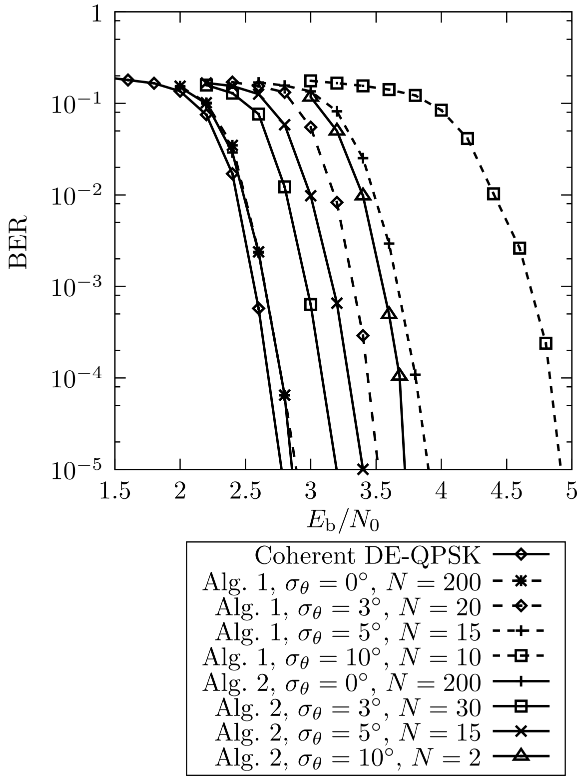

Fig. 4. BER performance of LDPC-coded DE-QPSK schemes where the proposed algorithms (both Algorithms 1 and 2) are used.

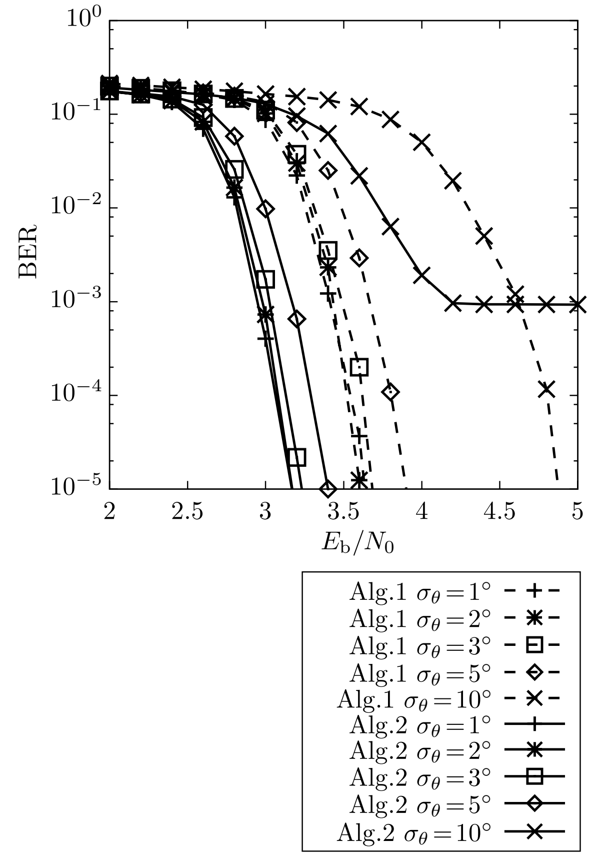

Fig. 5. BER performance, as a function of the SNR, of the proposed Algorithms 1 and 2. Several values of σ_θ are considered and $N = 15$.

are those of Fig. 4. The BER curves show clearly that for values of the phase noise parameter σ_θ lower than or equal to $5°$, decoding convergence is guaranteed for approximately the same SNR, whereas if $\sigma_\theta > 5°$ convergence is not guaranteed any longer, i.e., an error floor may appear. In particular, the error floor characterizing the BER curve corresponding to Algorithm 2 with $\sigma_\theta = 10°$ is due to the fact that, in order to cope with a strong phase noise, Algorithm 2 needs a very small inter-mix interval N, as clearly shown in Fig. 4. From the results in Fig. 5, one can conclude that the proposed algorithms are *blind* with respect to the phase noise intensity *as long as* this intensity is lower than that considered in the algorithm design.

In Fig. 6, the SNR needed to achieve a BER equal to 10^{-3} is shown as a function of σ_θ for both considered algorithms. The system and the simulation parameters are those of Fig. 4. One can conclude that the proposed algorithms are blind with respect to the phase noise intensity σ_θ, *as long as* this intensity is lower than a particular value which is a function of N. Beyond this critical value, the SNR needed to achieve the given BER value, i.e., 10^{-3}, diverges rapidly.

From the results in Figures 4, 5 and 6, one can conclude that, in the considered phase-uncertain channel scenario, Algorithm 2 performs better than Algorithm 1. This can be attributed to the strong approximations made by Algorithm 1 in "erasing" the phase memory at regular intervals, in an environment in which the phase varies slowly, yet continuously. This leads to wrong metrics in the proximity of mix epochs, where the erase operation is carried out. In other words, Algorithm 1 periodically enforces the strongest metrics among

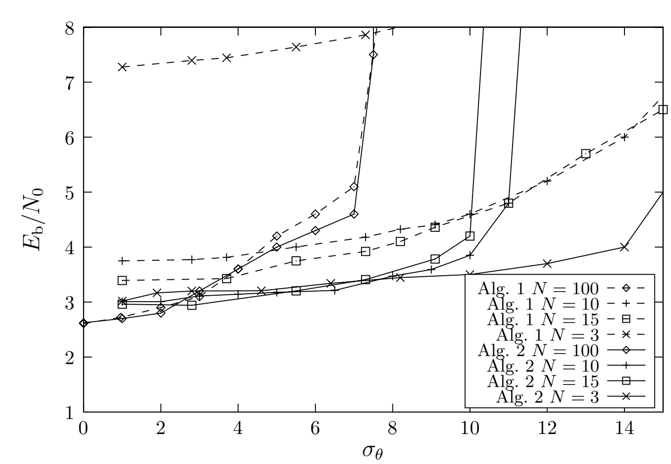

Fig. 6. SNR needed to achieve BER$= 10^{-3}$, as a function of σ_θ, considering Algorithms 1 and 2.

the trellises, whereas Algorithm 2 keeps the distribution of the metrics inside each trellis but periodically erases the different weightings of the trellises. This enables Algorithm 2 to account for channel phase variations during an N-symbol block.

IV. Detection by Multiple Trellises for Flat Fading Channels

In this section, the flat fading channel is considered. First, we derive the FB algorithm assuming a Markov chain model for the fading channel. Then, we specialize the algorithm introduced in Section II to the case of flat fading channel, highlighting its similarities with the Markov chain-based approach. Finally, the algorithms are analyzed and their performance is characterized through numerical results.

A. FB Algorithm for Fading Channels based on Markov Chains

The time-invariant flat fading model given in (5) can be extended to a more realistic model with time-varying flat fading. Accordingly, the discrete-time observable can be expressed as

$$r_k = f_k\, c_k + n_k \tag{12}$$

where $\{f_k\}$ is the fading process.[5] In the presence of Rayleigh fading, each fading realization f_k can be modeled as a zero-mean complex circularly symmetric Gaussian RV. We assume that the fading process $\{f_k\}$ is modeled according to Clarke [15], [16], with zero mean, unit variance and autocorrelation function $R_f(n) = J_0(2\pi n f_{\rm D} T)$, where $J_0(\cdot)$ is the zero-th order Bessel function and $f_{\rm D}T$ is the maximum normalized Doppler shift which characterizes the speed of the fading process.

We now outline the derivation of a simple first-order Markov chain model which approximately describes the evolution of the complex fading process. Several papers deal with Markov-chain modeling of the fading process—for more details, we refer the reader to [18], [29], [30] and references therein. We first partition the complex plane into $N_{\rm phase}$ angular sectors $[2\pi\frac{i-1}{N_{\rm phase}}, 2\pi\frac{i}{N_{\rm phase}})$, $i = 1, \ldots, N_{\rm phase}$. Then, we further split each sector into $N_{\rm ampl}$ "ring-shaped" regions. As a consequence, the complex plane is split into $N_{\rm phase}N_{\rm ampl}$ sub-domains $\{D_{ij}\}$ where D_{ij} denotes the domain corresponding to the i-th phase sector and the j-th ring-shaped region. In Fig. 7, an illustrative example with $N_{\rm phase} = 8$ angular sectors and $N_{\rm ampl} = 2$ ring-shaped regions is shown.

By associating the fading regions with states, it is possible to describe the evolution of the fading process through the use of a Markov chain. In general, considering a first-order Markov modeling for the fading process,[6] the total number of fading states is $L = N_{\rm ampl}N_{\rm phase}$. The probabilities of transition through different fading states can be computed through proper numerical integrations. For example, in order to evaluate the probability of transition from the region D_{ij} to the region D_{kl}, one can follow the method in [29], which is accurate as long as the first-order Markov chain modeling of the fading process holds and, in turns, corresponds to a scenario where the fading process is sufficiently *slow* [18].

Since the fading process is modeled through a Markov chain whose state corresponds to the current fading subregion D_{ij}, it is possible to derive a proper FB algorithm for the computation of the APPs of the transmitted symbols $\{a_k\}$. A general formulation accounting for a finite-memory channel depending

[5]We remark that this discrete-time model can be obtained from the continuous-time multiplicative fading model assuming that the fading process has a bandwidth much smaller than the signal bandwidth.

[6]We remark that the considered approach can easily be extended to higher-order Markov models of the fading process, at the expense of an increased number of fading states.

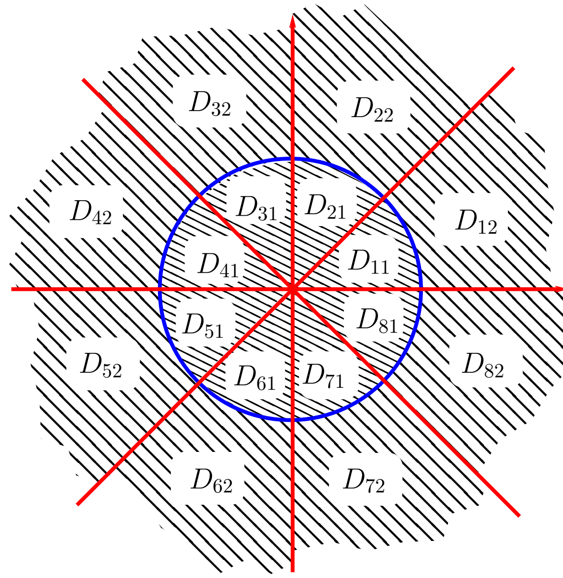

Fig. 7. Partitioning of the fading complex plane into fading regions.

on a generic process $\{\xi_k\}$ modeled by a Markov chain can be found in the Appendix. In particular, the FB algorithm in the Appendix assumes a channel whose output observables are independent conditionally on the data sequence and the parameter sequence. For instance, this is the case for the fading model in (12), where the observables are conditionally independent and Gaussian.

In the following, we will assume that the symbols $\{a_k\}$ are quaternary and encoded by a DE-QPSK encoder before transmission. The channel parameter ξ_k corresponds to the fading region $\hat{f}_k \in \{D_{ij}\} \quad i = 1, \ldots, N_{\rm phase},\ j = 1, \ldots, N_{\rm ampl}$. The extended state described in the Appendix here is $\sigma_k = (s_k, \tilde{f}_k)$, where s_k is the DE-QPSK encoder state at epoch k, and the fading region $\tilde{f}_k$ has been substituted to the generic parameter ξ_k.

The two essential ingredients needed for actual implementation of the Markov chain-based SISO algorithm in a scenario with fading are the transition probability $P\{\tilde{f}_{k+1}|\tilde{f}_k\}$ between the Markov chain states $\tilde{f}_k$ and $\tilde{f}_{k+1}$, obtained by suitably modeling the fading Markov chain, and the conditional PDF of the observable $p(r_k|a_k, \tilde{f}_k, s_k)$, given by the following

$$\begin{aligned} p(r_k|a_k, \tilde{f}_k, s_k) &= \frac{p(r_k, \tilde{f}_k|a_k, s_k)}{p\{\tilde{f}_k\}} \\ &= \frac{\displaystyle\int_{\tilde{f}_k} p(r_k|f, a_k, s_k) p_f(f) \mathrm{d}f}{\displaystyle\int_{\tilde{f}_k} p_f(f) \mathrm{d}f} \end{aligned} \tag{13}$$

where the independence between the fading process and the DE-QPSK coded data sequence c_k is exploited, $p(r_k|f, a_k, \sigma_k)$ is a Gaussian PDF (with mean fc_k), and $p_f(f)$ is the PDF of the fading coefficient.

B. Multi-Trellis SISO Algorithm for Fading Channel

The concept of detection by multiple trellises can be now directly applied to a fading channel. In particular, as for the phase-uncertain channel, if the channel is characterized by block-constant fading, Algorithm 1 is an optimum solution. In order to simplify the metric computation, the integral in (13) will be approximated by a finite sum of simple Gaussian metrics. We observed that this can lead to numerical problems at high SNR, where the noise variance becomes small. To overcome this problem, one may increase the accuracy of the numerical integration techniques used to compute (13), or prevent the variances of the Gaussian pdfs to become too small and trigger numerical problems.

Observe that every concatenated scheme with a powerful error correction code is characterized by a bad BER performance below a given SNR threshold and an operational BER performance beyond this threshold.[7] If the detection algorithm assumes a given, fixed, SNR value, one is guaranteed to obtain the performance of the same detection algorithm using the correct SNR value only when the actual SNR value and the fixed one are equal. Even with fixed SNR value assumed by the detection algorithm the BER as a function of the SNR is still expected to be monotonically decreasing. Therefore, if the assumed SNR is fixed to guarantee an operational BER at that very SNR value, the fixed SNR algorithm will guarantee operational BER beyond this SNR as well. As a consequence, we chose to fix the variance of the Gaussian metric, i.e., the SNR assumed by the detection algorithm, and to make it independent of the actual noise variance. This allows to overcome numerical problems and leads to a completely blind detection algorithm, which does not need either knowledge of fading or noise statistics.

C. Numerical Results

Unlike several works in the literature, where the fading process used in the simulations is generated according to the considered Markov-chain model, in the following the fading process used in the simulations is generated according to a realistic Clarke model.

In order to verify the effectiveness of the proposed detection by multiple trellises approach, we consider its application to the cases with uncoded DE-QPSK and with a regular (3,6) LDPC code with codeword length 32000—this length is necessary in order to combat long fades. The code should, in fact, "observe" a received sequence long enough to accurately describe the statistics of the channel, i.e., to exploit its ergodicity. We performed simulations considering $N_{\mathrm{ampl}} = 2$ and $N_{\mathrm{phase}} = 16$ and considering Algorithm 1 and the proposed simplified metric scheme. Algorithm 2, in the case of fading channel, exhibits unacceptable performance, and, therefore, is not shown. This is due to the fact that the mix operation in Algorithm 2 assigns large weights to trellises characterized by incorrect fading *amplitudes*. The considered normalized Doppler rate $f_{\mathrm{D}}T$ is equal to 0.01, corresponding to a moderately fast fading channel. The obtained results are shown in Fig. 8. The multi-trellis curve is obtained assuming a noise variance value corresponding to an SNR of about 7 dB. The inter-mix interval is heuristically optimized by trial

[7]In actual systems, the transition from bad BER performance to operational BER is not perfectly sharp, i.e., it happens within a small SNR region, usually referred to as *waterfall region*.

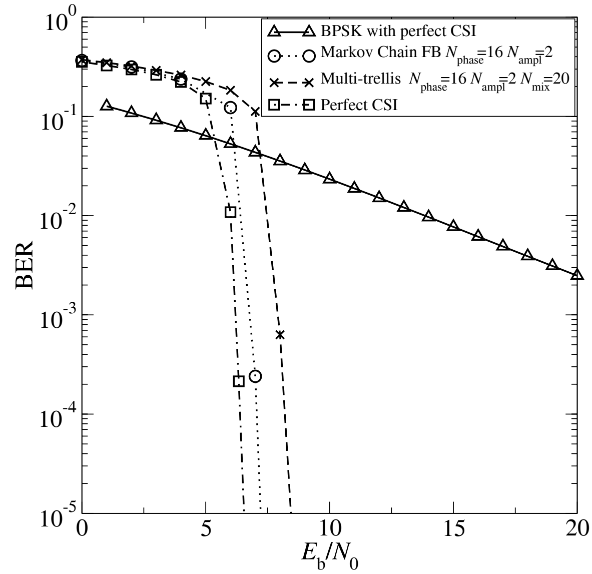

Fig. 8. BER performance, as a function of the SNR, in a scenario with a flat Rayleigh fading channel. Various schemes are considered: (i) BPSK with perfect CSI, (ii) LDPC-coded QPSK with Markov chain-based FB, (iii) an LDPC-coded QPSK with multi-trellis SISO, and (iv) LDPC-coded QPSK with perfect CSI.

and error and set to 20. In every LDPC-coded scheme, 30 decoding iterations are considered at the receiver side. The Markov chain-based algorithm presented in Section IV-A is also investigated and its performance is shown. As a reference, the performance of (i) the described concatenated scheme and (ii) an uncoded BPSK signaling, both considering perfect CSI, is also shown. As one can immediately see, the performance loss incurred by the use of the proposed detection by multiple trellises can be quantified at about 1 dB in comparison with the Markov-chain model performance and 1.8 dB compared with the perfect CSI scenario.

V. Complexity Analysis and Discussion

In this section, we investigate the complexity of the proposed multi-trellis SISO algorithms with a simple-minded, yet meaningful, approach. In order to highlight the advantages of the proposed algorithms, we compare their complexities with that of the general finite-memory approach to detection for channels affected by uncertain parameters described in [9] and with the Markov chain-based approach. We will evaluate the computational complexity in terms of elementary operations (i.e., additions or multiplications) *per trellis section* during a *single recursion*.

We preliminarily denote as $\mathrm{Comp}_{\mathrm{coher}}$ the complexity of an FB algorithm used by a coherent detector. It is possible to show that this complexity is

$$\mathrm{Comp}_{\mathrm{coher}} = \Theta(\Xi M)$$

where M is the cardinality of the information symbol set and the notation $\Theta(\cdot)$ stands for "on the order of." For simplicity, we assume that the complexity of the coherent receiver is the same in terms of multiplications and additions. Moreover, for ease of comparison between different algorithms, we assume that $\text{Comp}_{\text{coher}}$ is *exactly* ΞM.

We first evaluate the complexity of the proposed multi-trellis SISO algorithms, namely Algorithm 1, described in Subsection II-B1, and Algorithm 2, described in Subsection II-B2. For both algorithms, L trellis diagrams (each one equal to that of the coherent FB algorithm) are used. Therefore, this increases the complexity of the proposed SISO algorithms to $L\,\text{Comp}_{\text{coher}}$. At this point, one has to consider the additional complexity of the mix operations.

- For Algorithm 1, from the updating rule (7) one can conclude that 1 addition (over L quantized values of the channel parameter, i.e., phase or fading) for each state has to be carried out. Therefore, ΞL supplementary additions have to be considered. Since a mix operation takes place every N transitions, the complexity increase, in terms of elementary operations per trellis section, is $\Xi L/N$.
- For Algorithm 2, from the updating rule (8) one can conclude that 1 addition (over Ξ states) for each component trellis diagram has to be performed. At this point, one division has to be carried out per state and trellis component. Therefore, a mix operation requires, overall, $L\Xi$ additions and $L\Xi$ divisions. The complexity increase, per trellis section, is therefore, $2L\Xi/N$ in terms of elementary operations.

The finite-memory FB algorithm described in [9] is characterized by a "trellis expansion," in order to partially take into account the channel memory. This memory expansion is described by a finite-memory parameter N_{fm}, which characterizes the number of additional information symbols considered in the definition of a state in the trellis diagram at the receiver. The number of states for the computation of the state metrics in a detector/decoder, where a finite-memory FB algorithm is used, is $\Xi M^{N_{\text{fm}}}$. Therefore, one can conclude that the complexity increases proportionally to $M^{N_{\text{fm}}}$. Denoting by Comp_{fm} the complexity per trellis section (either in terms of additions or multiplications/divisions) in each recursion of a finite-memory FB algorithm, one can write:

$$\text{Comp}_{\text{fm}} = \text{Comp}_{\text{coher}} M^{N_{\text{fm}}} = \Xi M^{N_{\text{fm}}+1}.$$

The complexity of the first-order Markov chain model [18] analyzed in Section IV is proportional to the number of states of the coherent FB algorithm Ξ, to the cardinality of the input symbol space M, and to the square of the number of quantized values of the channel parameter, i.e., L^2. In other words, the complexity of a Markov chain-based FB algorithm is given by

$$\text{Comp}_{\text{MC}} = \text{Comp}_{\text{coher}} L^2 = \Xi M L^2.$$

The complexity of the proposed multi-trellis SISO algorithms, the finite-memory FB algorithm in [9], and the Markov chain-based FB algorithm are summarized in Table I. We remark that the complexity computation is, in practice, implementation-dependent. A noteworthy case is the implementation of the considered algorithms on a generic purpose processing unit, which usually leads to the serial computation of the quantities involved in the FB algorithms. In particular, it is well known that, due to the negative exponential behavior of the forward and backward state metrics, periodic normalization of these metrics, carried over all states and all trellises, is needed. This normalization takes place at arbitrary time epochs, but usually every 10–100 time steps. This normalization could be easily modified in order to implement the metric mixes described in this paper.

TABLE I
COMPLEXITY PER TRELLIS SECTION, DURING A SINGLE RECURSION, OF VARIOUS ALGORITHMS.

Algorithm	Complexity
SISO Algorithm 1	$\Xi M L + \frac{\Xi L}{N}$
SISO Algorithm 2	$\Xi M L + \frac{2\Xi L}{N}$
Finite-memory	$\Xi M^{N_{\text{fm}}+1}$
Markov Chain	$\Xi M L^2$

Moreover, the proposed multi-trellis SISO algorithms are *intrinsically* highly parallelizable. Exploiting properly this characteristic in the implementation could lead to significant increase of the decoding throughput, as desirable in future high data-rate wireless communication systems.

At this point, a careful reader can observe that since the finite-memory approach in [9] is very different from the multi-trellis SISO algorithms proposed in this paper, a meaningful complexity comparison between these algorithms should be carried out for a given performance level. To this purpose, we consider a complexity comparison for the same BER performance at the same SNR, using the same LDPC-coded DE-QPSK scheme used in Section III-B, in a scenario with phase noise. As shown in Fig. 9, the performance obtained by Algorithm 2 with $N = 15$ and $L = 8$, in a phase noise scenario with $\sigma_\theta = 5°$, is approximately approached by the finite-memory FB algorithm with $N_{\text{fm}} = 3$ (within a small fraction of a dB). The remaining simulation parameters are set as in Section III-B. The complexity of Algorithm 2 is $\Xi M L + 2L\Xi/N \simeq 32\Xi$, whereas the complexity of the finite-memory FB algorithm is $\Xi M^{3+1} = 256\Xi$. This large difference is due to the fact that the complexity of the finite-memory FB algorithm grows *exponentially* with the memory length (quantified by N_{fm}). This exponential increase of the complexity can be seen as a paradox characterizing the finite-memory approach, since slower phase processes require larger values of N_{fm}. As a consequence, good channels, exhibiting slow phase variations, need larger complexity than bad channels, with high phase dynamics. The proposed multi-trellis SISO algorithms are a pragmatic solution to overcome this paradox, since high values of N require lower complexity at the receiver.

In [8], [31], [32], other low complexity approaches to combat phase noise impairment are proposed, on the basis of a block-constant phase assumption. The additional strength of our multi-trellis SISO algorithms consists of their capability to take into account possible parameter changes within a block, in conjunction with the complete blindness with respect to the channel statistics.

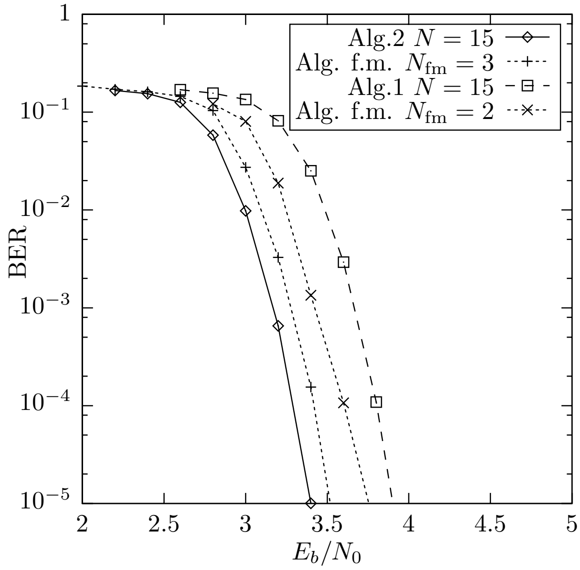

Fig. 9. BER performance of LDPC-coded DE-QPSK schemes where Algorithm 2 and a finite-memory (f.m.) detection algorithm are used. The phase noise parameter is $\sigma_\theta = 5°$.

VI. Conclusions

In this paper, we have introduced a novel approach, referred to as *detection by multiple trellises*, suitable for transmission over channels affected by time-varying parameters. After introducing the basic idea in an intuitive way, we have considered two relevant applications: detection for phase-uncertain and fading communications. The idea of the proposed approach consists of using several parallel trellises, over which *coherent* FB algorithms, each associated with a proper quantized value of the stochastic parameter, run. In order to cope with time-varying processes, the forward and backward metrics in the parallel FB algorithms are properly "mixed" together at regular intervals. For a scenario with phase noise, two multi-trellis SISO algorithms have been proposed, considering uncoded and LDPC-coded DE-QPSK transmission over an AWGN channel with Wiener phase noise. In the scenario with fading, after deriving an FB algorithm based on a simple first-order Markov chain model for the fading process, we have considered its multi-trellis extension. In all cases, DE-QPSK has been the used modulation format. An interesting feature of the proposed algorithms is the fact that they *do not* require knowledge of the statistics of the stochastic parameter (either the phase or fading), i.e., they are *blind*. Given their low complexity and high parallelizability, the proposed multi-trellis SISO algorithms are attractive for future high-throughput wireless communication systems. While the derivations have been carried out for FB algorithms [25], the proposed approach extends directly to trellis-based sequence detection algorithms, such as the Viterbi algorithm [33].

Appendix

In this appendix, an extension of the standard FB algorithm to a channel whose statistics at epoch k are a function of the state ξ_k of a Markov chain is described. Let us assume that, given $\{\xi_k\}$, the modulator-channel pair can be described by an FSM, in the sense that the observable statistics are functions of the state σ_k of an FSM whose input is the information symbol sequence $\{a_k\}$. Moreover, let us assume that (i) $\{a_k\}$ and $\{\xi_k\}$ are independent and (ii), given $\{a_k\}$ and $\{\xi_k\}$, the observables are independent. Following the guidelines in [9], [20], [21], it can be shown that the *a posteriori* probability of the symbol a_k can be computed as follows:

$$P\{a_k|\boldsymbol{r}_0^K\} = \sum_{(\sigma_k,\sigma_{k+1}):a_k} \beta_{k+1}(\sigma_{k+1})\,\alpha_k(\sigma_k)\,\gamma_k(\sigma_k,\sigma_{k+1},a_k) \tag{14}$$

where, as before, $\boldsymbol{r}_0^K$ denotes the vector of the observables and $\sigma_k = (s_k, \xi_k)$ is the (extended) state of the system; the notation $(\sigma_k, \sigma_{k+1}) : a_k$ denotes "the set of all (σ_k, σ_{k+1}) pairs compatible with the input symbol a_k" and the branch metric $\gamma_k(\sigma_k, \sigma_{k+1}, a_k)$ is defined as

$$\gamma_k(\sigma_k,\sigma_{k+1},a_k) = p(r_k|a_k,\xi_k,s_k)\cdot P\{a_k\}\cdot P\{\xi_{k+1}|\xi_k\} \tag{15}$$

in which $P\{\xi_{k+1}|\xi_k\}$ is the transition probability between the Markov chain states ξ_k and ξ_{k+1}, and $p(r_k|a_k, \xi_k, s_k)$ is the channel statistical description, i.e., the observable PDF given the data sequence and the channel parameter ξ_k. The forward and backward metrics $\alpha_k(\sigma_k)$ and $\beta_k(\sigma_k)$ are obtained with the following recursions:

$$\begin{aligned}\alpha_k(\sigma_k) &= \sum_{(\sigma_{k-1},a_{k-1}):\sigma_k} \alpha_{k-1}(\sigma_{k-1})\,\gamma_{k-1}(\sigma_{k-1},\sigma_k,a_{k-1})\\ \beta_k(\sigma_k) &= \sum_{(\sigma_{k+1},a_k):\sigma_k} \beta_{k+1}(\sigma_{k+1})\,\gamma_k(\sigma_k,\sigma_{k+1},a_k).\end{aligned}$$

The FB algorithm in (14) operates on a trellis whose number of states is the number Ξ of states of the modulator-channel FSM times the number L of states of the channel parameter Markov chain. This can be interpreted as a "super-trellis" comprising L trellises, each with Ξ states.

As special case, if the Markov chain $\{\xi_k\}$ is time-varying and the transition matrix differs from the identity matrix only at time epochs $k = Nl$, with $l \in \mathbb{N}$, it can be easily shown that the forward and backward recursions in the above extended FB algorithm are equivalent to the computation of L independent forward and backward recursions in the Ξ-state trellises for $N-1$ time steps. Every N time steps, the recursions involve, in general, all trellises. This corresponds to a block-constant discrete parameter ξ_k, which has been discussed in Section II-B assuming uniform distribution of the parameter realization. The corresponding super-trellis is shown in Fig. 1.

Acknowledgment

The authors wish to thank ing. Aldo Curtoni for his contributions to the simulation results.

References

[1] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261-1271, Oct. 1996.
[2] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 22-24, Jan. 1997.
[3] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers (Applications of Communications Theory).* Plenum Press, 1997.

[4] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[5] A. Anastasopoulos and K. M. Chugg, "Adaptive soft-input soft-output algorithms for iterative detection with parametric uncertainty," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1638-1649, Oct. 2000.
[6] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1376-1385, Sept. 1999.
[7] G. Ferrari, G. Colavolpe, and R. Raheli, "Linear predictive receivers for phase uncertain channels," in *Proc. Intern. Symp. Signal Process. Applications*, vol. 1, Paris, France, July 2003, pp. 301-304.
[8] C. Rong-Rong, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: a practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1676-1689, Oct. 2003.
[9] G. Ferrari, G. Colavolpe, and R. Raheli, "A unified framework for finite-memory detection," *IEEE J. Select. Areas Commun.*, vol. 23, no. 9, pp. 1697-1706, Sept. 2005.
[10] 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, June 1990.
[11] 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.
[12] 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.
[13] 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.
[14] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol MAP demodulation of CPM and PSK signals on Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 788-799, July 1997.
[15] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell System Tech. J.*, vol. 47, pp. 957-1000, Aug. 1968.
[16] W. C. Jakes, *Microwave Mobile Communications*. New York: John Wiley & Sons, 1974.
[17] H. S. Wang and N. Moayeri, "Finite-state Markov channel–a useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163-171, Feb. 1995.
[18] C. C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2032-2040, Dec. 2000.
[19] W. Turin, R. Jana, C. Martin, and J. Winters, "Modeling wireless channel fading," in *Proc. IEEE Veh. Tech. Conf.* (VTC), pp. 1740-1744, Oct. 2001.
[20] O. Macchi and L. Scharf, "A dynamic programming algorithm for phase estimation and data decoding on random phase channels," *IEEE Trans. Inform. Theory*, no. 5, pp. 581-595, Sept. 1981.
[21] M. Peleg, S. Shamai (Shitz), and S. Galán, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel," in *IEE Proceedings-Commun.*, vol. 147, no. 2, pp. 87-95, Apr. 2000.
[22] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1706-1717, Sept. 2001.
[23] 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.
[24] M. Riediger and E. Shwedyk, "Communication receivers based on Markov models of the fading channel," *IEE Proc.-Commun.*, vol. 150, no. 4, pp. 275-279, Aug. 2003.
[25] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284-287, Mar. 1974.
[26] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
[27] T. Keller and L. Hanzo, "Adaptive modulation techniques for duplex OFDM transmission," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1893-1906, Sept. 2000.
[28] M. Franceschini, G. Ferrari, R. Raheli, and A. Curtoni, "Serial concatenation of LDPC codes and differential modulations," *IEEE J. Select. Areas Commun.*, vol. 23, no. 9, pp. 1758-1768, Sept. 2005.
[29] W. Turin and R. V. Nobelen, "Hidden Markov modeling of flat fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1809-1817, Dec. 1998.
[30] F. Babich, O. E. Kelly, and G. Lombardi, "Generalized Markov modeling for flat fading," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 547-551, Apr. 2000.
[31] J. K. M. Mackenthun, "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 33, no. 2-4, pp. 1471-1474, Feb./Mar./Apr. 1994.
[32] I. Motedayen-Aval and A. Anastasopoulos, "Polynomial-complexity noncoherent symbol-by-symbol detection with application to adaptive iterative decoding of turbo-like codes," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 197-207, Feb. 2003.
[33] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268-278, Mar. 1973.

Michele Franceschini was born in Milan, Italy, in 1977. He received the Dr. Ing. degree ("Laurea," 5-year program) in Electrical Engineering (summa cum laude) from the University of Parma in 2002. In 2003, he received the "Paolo Conti Award" as the best graduate in Information Engineering at the University of Parma in the academic year 2002. In 2006, he received the Ph.D. from the University of Parma. From 2006 to 2008 he held a postdoc position at the University of Parma. Since March 2008, he holds a postdoc position at the IBM T.J. Watson Research Center, Yorktown Heights, NY. His research interests lie in the area of communication and information theory, storage channels, low-density parity-check code design, theoretical aspects of optical communication, advanced signal processing techniques, synchronization, and low-complexity implementation of digital communication systems.

Gianluigi Ferrari was born in Parma, Italy, in November 1974. He received the "Laurea" degree (five-year program) (summa cum laude) in electrical engineering and the Ph.D. degree in Information Technologies from the University of Parma in October 1998 and January 2002, respectively. From July 2000 to December 2001, he was a Visiting Scholar at the Communication Sciences Institute, University of Southern California, Los Angeles, CA, USA. Since 2002, he has been a Research Professor with the Department of Information Engineering, University of Parma, where he is now the coordinator of the Wireless Ad-hoc and Sensor Networks (WASN) Laboratory. Between 2002 and 2004, he visited several times, as a Research Associate, the Electrical and Computer Engineering Department at Carnegie Mellon University, Pittsburgh, PA. In fall 2007 he visited, as a DUO-Thailand Fellow, the King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand.

Dr. Ferrari has published more than 100 papers in leading international conferences and journals. He is coauthor of the books *Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems* (Wiley: 2004), *Introduzione a Teoria della probabilità e variabili aleatorie con applicazioni all'ingegneria e alle scienze*, (Editrice Esculapio-Progetto Leonardo:2008), *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective* (Wiley: 2006), and *LDPC Coded Modulations* (Springer: 2009). His research interests include digital communication systems analysis and design, wireless ad hoc and sensor networking, adaptive digital signal processing, and information theory.

Dr. Ferrari is a co-recipient of a best student paper award at the 2006 International Workshop on Wireless Ad hoc Networks (IWWAN'06). He acts as a frequent reviewer for many international journals and conferences. He acts also as a technical program member for many international conferences. He currently serves on the Editorial Boards of THE OPEN ELECTRICAL AND ELECTRONIC ENGINEERING (TOEEJ) JOURNAL (Bentham Publishers), the INTERNATIONAL JOURNAL OF RF TECHNOLOGIES: RESEARCH AND APPLICATIONS (Taylor & Francis), and the INTERNATIONAL JOURNAL OF FUTURE GENERATION COMMUNICATION AND NETWORKING (SERSC: Science & Engineering Research Support Center).

Riccardo Raheli received the Dr. Ing. (Laurea) degree in Electrical Engineering "summa cum laude" from the University of Pisa, Italy, in 1983, the Master of Science (M.S.) degree in Electrical and Computer Engineering from the University of Massachusetts at Amherst, U.S.A., in 1986, and the Doctoral (Perfezionamento) degree in Electrical Engineering "summa cum laude" from the Scuola Superiore S. Anna di Studi Universitari e di Perfezionamento, Pisa, Italy, in 1987.

From 1986 to 1988 he was with Siemens Telecomunicazioni, Milan, Italy. From 1988 to 1991, he was a Research Professor at the Scuola Superiore S. Anna, Pisa, Italy. In 1990, he was a Visiting Assistant Professor at the University of Southern California, Los Angeles, U.S.A.. Since 1991, he has been with the University of Parma, Italy, first as a Research Professor, then Associate Professor and currently Professor of Communications Engineering and Chairman of the Communications Engineering Program Committee. 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 work has lead to numerous scientific publications in leading international journals and conference proceedings, as well as a few industrial patents. In 1990, he conceived (with Andreas Polydoros) the principle of "Per-Survivor Processing." He is coauthor of a few scientific monographs such as *Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems* (John Wiley & Sons, 2004) and *LDPC Coded Modulations* (Springer, 2009). He is coauthor of the paper which received the "2006 Best Student Paper Award in Signal Processing & Coding for Data Storage" from the IEEE Communications Society.

Dr. Raheli served on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS, as an Editor for Detection, Equalization and Coding, from 1999 to 2003. He was the leading Guest Editor of a special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) on "Differential and Noncoherent Wireless Communications," published in 2005. He served on the Editorial Board of the "European Transactions on Telecommunications (ETT)," as an Editor for Communication Theory, from 2003 to 2008. He has also served on the Technical Program Committee of many leading international conferences in the areas of Communication Theory and Systems, Information Theory and Signal Processing, such as the IEEE International Conference on Communications (ICC), IEEE Global Communications Conference (GLOBECOM), European Signal Processing Conference (EUSIPCO), IEEE International Symposium on Power-Line Communications and Its Applications (ISPLC), International Symposium on Information Theory and its Applications (ISITA), International Symposium on Spread Spectrum Techniques and Applications (ISSSTA), and others.