Soft-Output Decoding of Rotationally Invariant Codes Over Channels With Phase Noise

Alan Barbieri and Giulio Colavolpe, *Member, IEEE*

***Abstract*—We consider rotationally invariant (RI) trellis-coded modulations (TCMs) transmitted over channels affected by phase noise. To describe the main ideas of this paper, we first concentrate, as a case study, on the simplest RI scheme, namely the differentially encoded M-ary phase-shift keying (M-PSK) signal. For this problem, we use the framework based on factor graphs (FGs) and the sum–product algorithm (SPA), to derive the *exact* maximum *a posteriori* (MAP) symbol detection algorithm. By analyzing its properties, we demonstrate that it can be implemented by a forward–backward estimator of the phase probability density function, followed by a symbol-by-symbol completion to produce the *a posteriori* probabilities of the information symbols. To practically implement the forward–backward phase estimator, we propose a couple of schemes with different complexity. The resulting algorithms exhibit an excellent performance and, in one case, only a limited complexity increases with respect to the algorithm that perfectly knows the channel phase. The properties of the optimal decoder and the proposed practical decoding schemes are then extended to the case of a generic RI code. The proposed soft-output algorithms can also be used in iterative decoding schemes for concatenated codes employing RI inner components. Among them, in the numerical results, we consider repeat-accumulate (RA) codes and other serially concatenated schemes recently proposed in the technical literature.**

***Index Terms*—Differential encoding, maximum *a posteriori* (MAP) symbol detection, repeat-accumulate codes, rotationally invariant codes.**

I. Introduction

IN THE last few years, the problem of robust decoding in channels affected by a time-varying phase has been investigated for classical coding schemes (see [1] and references therein), as well as for powerful channel codes to be decoded iteratively, such as turbo codes or low-density parity-check codes [2]–[10]. In particular, the authors in [10] developed an algorithm with a very low complexity and a practically optimal performance. Some of these algorithms require the insertion of pilot symbols in order to solve the phase ambiguity problem that arises in phase-uncertain channels, and to make the iterative decoder bootstrap, especially in the case of strong phase noise and long codeword lengths [10].

Paper approved by A. Anastasopoulos, the Editor for Iterative Detection, Estimation and Coding of the IEEE Communications Society. Manuscript received April 26, 2006; revised November 7, 2006. This work was supported by the European Space Agency (ESA)-European Space and Research Technology Centre (ESTEC), Noordwijk, The Netherlands, under contract no. 19370/05/NL/JD. The paper was presented in part at the IEEE International Symposium on Information Theory (ISIT'06), Seattle, WA, July 2006, and at the 14th European Signal Processing Conference (EUSIPCO'06), Florence, Italy, September 2006.

A. Barbieri is with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy, and is also with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: barbieri@tlc.unipr.it).

G. Colavolpe is with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy (e-mail: giulio@unipr.it).

Digital Object Identifier 10.1109/TCOMM.2007.908520

An alternative to pilot symbols, that avoids the decrease of the effective information rate due to pilot insertion, is represented by the use of a rotationally invariant (RI) code [11]–[14], such as a differential code [15]. An inner differential encoder is used as an inner component in most of the serially concatenated RI schemes recently proposed in the literature [16]–[19]. By using an equivalent terminology, an inner accumulator, is also a component of repeat-accumulate (RA) codes [20], [21].

In principle, the algorithms described in [6]–[10] can be used for RI coding schemes. However, they have a main drawback: they perform separate detection and phase tracking, namely, at every iteration an instance of a soft-input soft-output (SISO) decoding algorithm for the RI code, for example implemented by means of a Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [22], the execution of the code-aware phase tracking algorithm (which takes advantage from the a posteriori probabilities coming from the BCJR), and finally, another execution of the BCJR, are performed. Hence, they are characterized by a higher latency, and do not exploit the code structure but only the soft outputs produced by the decoder, thus requiring the insertion of (a minimal amount of) pilot symbols to bootstrap.

On the contrary, the algorithms in [2]–[5] can be designed to jointly perform the decoding of the RI and the detection in the presence of the unknown time-varying phase. In [2], after a proper discretization of the phase space, a super-trellis, taking into account the RI code trellis and the phase model, is built and the BCJR algorithm is run over it. In [3] and [4], the channel phase is *a priori* averaged out, but the resulting algorithm still works on an expanded trellis. Finally, the algorithm in [5] can work on the trellis of the RI encoder or on an expanded trellis, and *multiple* nonBayesian phase estimators are used in the forward and backward recursions of the algorithm.

In this paper, to illustrate the main concepts behind the derivation of the proposed algorithms, we first consider the problem of a differentially encoded M-ary phase-shift keying (M-PSK) signal transmitted over a channel affected by phase noise. The approach is Bayesian, i.e., the channel phase is modeled as a stochastic process with known statistics. Although the implementation of the exact maximum *a posteriori* (MAP) symbol detection algorithm is impractical, we analyze its properties, finding that it can be implemented by using a *single* forward–backward estimator of the phase probability density function (pdf), followed by a symbol-by-symbol completion to produce the *a posteriori* probabilities of the information symbols. This algorithm obviously works in a noniterative joint decoding/phase tracking fashion, and does not require the insertion of pilot symbols. Then, by using the canonical distribution approach [23], we develop a couple of practical schemes to implement the forward–backward estimator.

The resulting algorithms may be used as SISO blocks for iterative detection/decoding in concatenated schemes using an inner differential encoder. We also describe the extension of the proposed technique to generic RI codes and show the relevant numerical results. Although in the numerical results, we mainly concentrate on iteratively decodable concatenated codes, since the problem of detection in the presence of phase noise is made harder by the low operating SNR, the proposed schemes can also be adopted for the noniterative joint detection and decoding of a single (nonconcatenated) RI code.

The remainder of this paper is organized as follows. In Section II, we provide the system model. By means of the framework based on factor graphs (FGs) and the sum–product algorithm (SPA) [24], in Section III, the exact MAP symbol detection algorithm is derived and its properties analyzed. The low-complexity algorithms, based on the approximated versions of the exact detection strategy, are described in Section IV. The extension to RI codes is discussed in Section V. The performance of the proposed schemes, obtained through computer simulations, is assessed in Section VI, and finally, in Section VII, some conclusions are drawn.

II. System Model

We consider the transmission of a sequence of $K+1$ complex modulation symbols $\mathbf{c} = \{c_k\}_{k=0}^{K}$, belonging to an M-PSK alphabet $\{e^{j2\pi/M\, i}, i = 0, 1, \ldots, M-1\}$, over an additive white Gaussian noise (AWGN) channel affected by an unknown time-varying phase. Symbols $\{c_k\}$ are obtained from information symbols $\{a_k\}_{k=1}^{K}$, assumed independent, but not identically nor uniformly distributed, and belonging to the same M-PSK alphabet, through differential encoding, i.e.,

$$c_k = c_{k-1} a_k\ . \tag{1}$$

Symbol c_{k-1} is also the encoder state. Since the transmission over a channel affected by phase noise will be considered, we may assume that the initial symbol c_0 is unknown to the receiver due to the initial channel phase uncertainty. Assuming Nyquist transmitted pulses, matched filtering, phase variations to be slow enough so that no intersymbol interference arises, the discrete-time baseband received signal is given by

$$r_k = c_k e^{j\theta_k} + w_k, \qquad k = 0, 1, \ldots, K \tag{2}$$

where the noise samples $\mathbf{w} = \{w_k\}_{k=0}^{K}$ are independent and identically distributed (i.i.d.), complex, circularly symmetric Gaussian random variables, each with zero mean and variance $2\sigma^2$, σ^2 being the variance per component.

In the derivation of the proposed algorithms, for the time-varying channel phase θ_k we assume a random-walk (Wiener) model $\theta_{k+1} = \theta_k + \Delta_k$, where $\{\Delta_k\}$ are real i.i.d. Gaussian random variables with zero mean and standard deviation σ_Δ,[1] and θ_0 is uniformly distributed in $[0, 2\pi)$. In the derivation of the proposed algorithms, the value of σ_Δ is assumed to be known to the receiver. The sequence of phase increments $\{\Delta_k\}$ is supposed to be unknown to both the transmitter and the receiver, and statistically independent of $\mathbf{c}$ and $\mathbf{w}$. The assumption on the phase noise model will be relaxed in the numerical results.

III. MAP Symbol Detection of Differentially Encoded PSK Signals

Here, we derive the exact MAP symbol detection algorithm for the considered problem by using a properly defined FG and the SPA [24].

Let us consider the joint distribution of vectors $\mathbf{a}$, $\mathbf{c}$, and $\boldsymbol{\theta} = \{\theta_k\}_{k=0}^{K}$, given $\mathbf{r} = \{r_k\}_{k=0}^{K}$:[2]

$$\begin{aligned} p(\mathbf{a}, \mathbf{c}, \boldsymbol{\theta} | \mathbf{r}) &\propto P(\mathbf{a}) P(\mathbf{c}|\mathbf{a}) p(\boldsymbol{\theta}) p(\mathbf{r}|\mathbf{c}, \boldsymbol{\theta}) \\ &= P(\mathbf{a}) P(\mathbf{c}|\mathbf{a}) p(\boldsymbol{\theta}) \prod_{k=0}^{K} p(r_k | c_k, \theta_k) \end{aligned} \tag{3}$$

where

$$p(r_k | c_k, \theta_k) = \frac{1}{2\pi\sigma^2} \exp\left\{ -\frac{|r_k - c_k e^{j\theta_k}|^2}{2\sigma^2} \right\}. \tag{4}$$

We can further factor the terms $P(\mathbf{a})$, $P(\mathbf{c}|\mathbf{a})$, and $p(\boldsymbol{\theta})$ in (3) as

$$P(\mathbf{a}) = \prod_{k=1}^{K} P(a_k) \tag{5}$$

$$P(\mathbf{c}|\mathbf{a}) = P(c_0) \prod_{k=1}^{K} I(c_k, c_{k-1}, a_k) \tag{6}$$

$$p(\boldsymbol{\theta}) = p(\theta_0) \prod_{k=1}^{K} p(\theta_k | \theta_{k-1}) \tag{7}$$

where $I(c_k, c_{k-1}, a_k)$ is an indicator function, equal to 1 if a_k and the differential symbols c_k and c_{k-1} respect the constraint (1), and zero otherwise. Since the SPA is defined up to scaling its messages by positive factors, independent of the variables represented in the graph, from now on, we take the notational liberty of using the equality symbol "$=$" instead of the proportionality symbol "$\propto$." Substituting (5)–(7), into (3), *clustering* [24] the variables c_k and θ_k, we obtain the FG in Fig. 1. Since this FG does not contain cycles, the application of the SPA to it, with

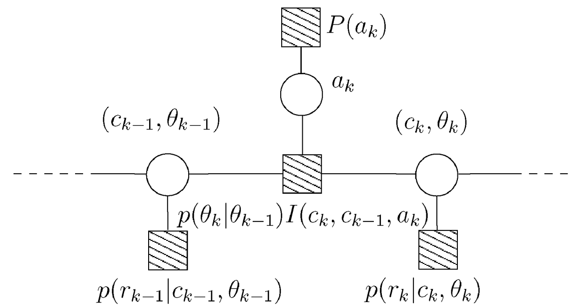

Fig. 1. Factor graph for the considered problem.

[1] Note that, since the channel phase is defined modulo 2π, the pdf $p(\theta_{k+1}|\theta_k)$ can be approximated as Gaussian in θ_{k+1}, with mean θ_k and variance σ_Δ^2, only if $\sigma_\Delta \ll 2\pi$.

[2] We still use the symbol $p(\cdot)$ to denote a continuous pdf with some discrete probability masses.

a *noniterative* forward–backward schedule, produces the *exact a posteriori* probabilities of the symbols $\{a_k\}$. Taking into account the probabilistic meaning of the messages in the graph, defining $\mathbf{r}_{k_1}^{k_2} = \{r_k\}_{k=k_1}^{k_2}$ (hence $\mathbf{r} = \mathbf{r}_0^K$), the forward and backward recursions and the completion necessary to compute the *a posteriori* probabilities of the symbol a_k (or, equivalently, the extrinsic information $P(a_k|\mathbf{r})/P(a_k)$) are, respectively,

$$p(c_k, \theta_k|\mathbf{r}_0^k) = p(r_k|c_k, \theta_k) \sum_{a_k} P(a_k) \cdot \int p(c_{k-1} = c_k a_k^*, \theta_{k-1}|\mathbf{r}_0^{k-1}) p(\theta_k|\theta_{k-1}) d\theta_{k-1}$$

$$p(c_{k-1}, \theta_{k-1}|\mathbf{r}_{k-1}^K) = p(r_{k-1}|c_{k-1}, \theta_{k-1}) \sum_{a_k} P(a_k) \cdot \int p(c_k = c_{k-1} a_k, \theta_k|\mathbf{r}_k^K) p(\theta_k|\theta_{k-1}) d\theta_k \quad (8)$$

$$\frac{P(a_k|\mathbf{r})}{P(a_k)} = \sum_{c_{k-1}} \iint p(c_{k-1}, \theta_{k-1}|\mathbf{r}_0^{k-1}) \quad (9)$$

$$p(c_k = c_{k-1} a_k, \theta_k|\mathbf{r}_k^K) p(\theta_k|\theta_{k-1}) d\theta_{k-1} d\theta_k. \quad (10)$$

The initializations for the forward and backward recursions are $p(c_0, \theta_0|r_0) = p(r_0|c_0, \theta_0)$ and $p(c_K, \theta_K|r_K) = p(r_K|c_K, \theta_K)$

Equations (8)–(10) can be equivalently obtained by using a (more involved) probabilistic derivation. In fact, the received signal is the output of a Markov source observed through an AWGN channel. As a consequence, this detection problem is the same as in [22], with the only difference that the "state" of the Markov source is defined, in this case, as the joint mixed discrete-continuous random variable (c_{k-1}, θ_{k-1}). Hence, a forward-backward algorithm similar to that in [22] results.

Let us now consider (8). We may decompose

$$p(c_k, \theta_k|\mathbf{r}_0^k) = p(\theta_k|c_k, \mathbf{r}_0^k)\; P(c_k|\mathbf{r}_0^k). \quad (11)$$

The first term of the right-hand side considers the distribution of the unknown phase, at time k, given the past and present received samples, and the state of the differential encoder, while the second term is the state probability that is exactly the same probability mass function (pmf) evaluated in case of detection in the presence of a known phase. In practice, the algorithm performs a *per-state Bayesian estimation* of the channel phase during the forward recursion. A similar decomposition can be clearly accomplished for the backward pdf (9).

A proof of the following three properties is given in the Appendix.

Property 1: Irrespective of the values of the *a priori* information $\{P(a_k)\}$, the state probabilities are $P(c_k|\mathbf{r}_0^k) = \text{const.}$ and $P(c_k|\mathbf{r}_k^K) = \text{const.}$, for each value of k. Hence, it is not necessary to evaluate them.

Property 2: The pdfs $p(\theta_k|c_k, \mathbf{r}_0^k)$, for different values of c_k, differ for a shift of a multiple of $2\pi/M$, i.e.,

$$p(\theta_k|c_k = e^{j\frac{2\pi}{M}i}, \mathbf{r}_0^k) = p\left(\theta_k + \frac{2\pi}{M}i|c_k = e^{j0}, \mathbf{r}_0^k\right). \quad (12)$$

An identical result holds for the backward pdf $p(\theta_k|c_k, \mathbf{r}_k^K)$.

Property 3: The summation over c_{k-1} in the completion (10) disappears because all the M terms of the summation are equal. Hence, only one of them needs to be evaluated.

From these three properties, it follows that, for each time epoch k, only the pdfs $p(\theta_k|c_k = 1, \mathbf{r}_0^k)$ and $p(\theta_k|c_k = 1, \mathbf{r}_k^K)$ need to be evaluated. By defining $\alpha_k(\theta_k) \triangleq p(\theta_k|c_k = 1, \mathbf{r}_0^k)$ and $\beta_k(\theta_k) \triangleq p(\theta_k|c_k = 1, \mathbf{r}_k^K)$, the forward-backward algorithm described by (8)–(10) simplifies to

$$\alpha_k(\theta_k) = p(r_k|c_k = 1, \theta_k) \sum_{i=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) \cdot \int \alpha_{k-1}\left(\theta_{k-1} - \frac{2\pi}{M}i\right) p(\theta_k|\theta_{k-1}) d\theta_{k-1} \quad (13)$$

$$\beta_{k-1}(\theta_{k-1}) = p(r_{k-1}|c_{k-1} = 1, \theta_{k-1}) \sum_{i=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) \cdot \int \beta_k\left(\theta_k + \frac{2\pi}{M}i\right) p(\theta_k|\theta_{k-1}) d\theta_k \quad (14)$$

$$\frac{P(a_k = e^{j\frac{2\pi}{M}i}|\mathbf{r})}{P(a_k = e^{j\frac{2\pi}{M}i})} = \iint \alpha_{k-1}(\theta_{k-1}) \beta_k\left(\theta_k + \frac{2\pi}{M}i\right) \cdot p(\theta_k|\theta_{k-1}) d\theta_{k-1} d\theta_k. \quad (15)$$

Hence, we have a single forward–backward estimator of the phase pdf and a final completion.

This exact MAP symbol detection strategy involves integration and computation of continuous pdfs, and it is not suited for direct implementation. A solution for this problem is suggested in [23] and consists of the use of *canonical distributions*, i.e., the pdfs $\alpha_k(\theta_k)$ and $\beta_k(\theta_k)$ computed by the algorithm are constrained to be in a certain "canonical" family, characterized by some parameterization. Hence, the forward and backward recursions reduce to propagating and updating the parameters of the pdf rather than the pdf itself. In the next section, two algorithms based on this approach will be described.

IV. Low-Complexity Algorithms

A. First Algorithm

A very straightforward solution to implement (13) and (14) is obtained by discretizing the channel phase [2], [10]. This approach is also similar to that proposed in [25], for flat fading channels. In this way, the pdfs $\alpha_k(\theta_k)$ and $\beta_k(\theta_k)$ become pmfs and the integrals in (13)–(15) become summations. When the number L of discretization levels is large enough, at least $L = 8M$ [2], the resulting algorithm becomes optimal (in the sense that its performance approaches that of the exact algorithm). Hence, it may also be used to obtain a performance benchmark and will be denoted as "discretized-phase algorithm" (*dp-algorithm*). Note that with respect to the algorithm proposed in [2], we are exploiting here the properties in Section III to reduce the algorithm complexity. The performance of the described *dp-algorithm* is, therefore, identical for a given value of L, to that of the algorithm proposed in [2], but with a lower complexity.

B. Second Algorithm

By observing that the Tikhonov distribution ensures a very interesting performance with a low complexity when used as a canonical distribution in detection algorithms for phase noise channels, as demonstrated in [10], and observing the form of the pmfs obtained by adopting the described *dp-algorithm*, pdfs $\alpha_k(\theta_k)$ and $\beta_k(\theta_k)$ are constrained to have the following expressions

$$\alpha_k(\theta_k) = \sum_{m=0}^{M-1} q_{f,k}^{(m)} t\left(z_{f,k}\, e^{j\frac{2\pi}{M}m}; \theta_k\right) \qquad (16)$$

$$\beta_k(\theta_k) = \sum_{m=0}^{M-1} q_{b,k}^{(m)} t\left(z_{b,k}\, e^{j\frac{2\pi}{M}m}; \theta_k\right) \qquad (17)$$

where, for each time index k, $\{q_{f,k}^{(m)}, m = 0, 1, \ldots, M-1\}$ ($\{q_{b,k}^{(m)}, m = 0, 1, \ldots, M-1\}$) and $z_{f,k}$ ($z_{b,k}$) are, respectively, M real coefficients and one complex coefficient, and $t(z;\theta)$ is the Tikhonov distribution with complex parameter z defined as

$$t(z;\theta) = \frac{e^{\mathrm{Re}\left[z\, e^{-j\theta}\right]}}{2\pi \mathrm{I}_0(|z|)} \qquad (18)$$

$\mathrm{I}_0(x)$ being the zeroth-order modified Bessel function of the first kind.

Three approximations are now introduced in order to derive a low complexity detection algorithm. A justification of these approximations is represented by the excellent performance of the resulting algorithm.

1) the convolution of a Tikhonov and a Gaussian pdf is still a Tikhonov pdf, with a modified complex parameter [10], [26], i.e.,

$$\int t(z;x) g(x, \rho^2; y) dx \simeq t\left(\frac{z}{1+\rho^2|z|}; y\right) \qquad (19)$$

where $g(x, \rho^2; y)$ represents a Gaussian pdf in y with mean x and variance ρ^2

2) Since, for large arguments, $\mathrm{I}_0(x) \simeq e^x$, we approximate

$$e^{\mathrm{Re}[z\, e^{-j\theta}]} \simeq 2\pi e^{|z|} t(z;\theta). \qquad (20)$$

3) Let z be a complex number, $\{u_m, m = 0, 1, \ldots, M-1\}$ a set of complex numbers, and $\{q_m, m = 0, 1, \ldots, M-1\}$ a set of real numbers such that $\sum_m q_m = 1$, then the following approximation holds, especially when $|z|$ is sufficiently larger than each $|u_m|$ or when there is an $\overline{m}$ such that $q_{\overline{m}} \gg q_m$, $\forall m \neq \overline{m}$:

$$\sum_m q_m t\left(z\, e^{j\frac{2\pi}{M}m} + u_m; \theta\right) \simeq \sum_m q_m t\left(w\, e^{j\frac{2\pi}{M}m}; \theta\right) \qquad (21)$$

where $w = z + \sum_\ell q_\ell u_\ell e^{-j\frac{2\pi}{M}\ell}$.

We now derive the reduced-complexity forward recursion. Substituting (4) into (13), assuming that $\alpha_{k-1}(\theta_{k-1})$ has the canonical expression (16), and using approximation (19), we obtain

$$\alpha_k(\theta_k) = e^{\frac{1}{\sigma^2}\mathrm{Re}\left[r_k e^{-j\theta_k}\right]} \sum_{i=0}^{M-1}\sum_{m=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) \cdot q_{f,k-1}^{(m)} t\left(z'_{f,k-1}; \theta_k - \frac{2\pi}{M}(m+i)\right) \qquad (22)$$

where $z'_{f,k-1} = z_{f,k-1}/(1+\sigma_\Delta^2|z_{f,k-1}|)$. Now, by changing the first summation index in $n = (i+m)_{\mathrm{mod}\,M}$, using (18) and (20), and discarding irrelevant multiplicative factors, we have

$$\alpha_k(\theta_k) = \sum_{n=0}^{M-1}\left[\sum_{i=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) q_{f,k-1}^{(n-i)_{\mathrm{mod}\,M}}\right] \cdot e^{\left|z'_{f,k-1}e^{j\frac{2\pi}{M}n} + \frac{r_k}{\sigma^2}\right|} t\left(z'_{f,k-1}e^{j\frac{2\pi}{M}n} + \frac{r_k}{\sigma^2}; \theta_k\right). \qquad (23)$$

This resulting $\alpha_k(\theta_k)$ is not in the constrained form (16). However, by applying the approximation (24), we obtain the following updating equations for the parameters of the canonical distribution (16)

$$q_{f,k}^{(m)} \propto \left[\sum_{i=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) q_{f,k-1}^{(m-i)_{\mathrm{mod}\,M}}\right], \cdot e^{\left|z'_{f,k-1}e^{j\frac{2\pi}{M}m} + \frac{r_k}{\sigma^2}\right|}, \qquad m = 0, \ldots, M-1 \qquad (24)$$

$$z_{f,k} = z'_{f,k-1} + \frac{r_k}{\sigma^2}\sum_{m=0}^{M-1} q_{f,k}^{(m)}\, e^{-j\frac{2\pi}{M}m}. \qquad (25)$$

It is worth noticing that, before the evaluation of the coefficient $z_{f,k}$, the M real coefficients $q_{f,k}^{(m)}$ evaluated through (24) have to be normalized so that their sum is 1. Since there is no *a priori* knowledge of the initial phase or of the initial differential symbol, the following initial values of the recursive coefficients result

$$q_{f,0}^{(m)} = \delta_m \qquad z_{f,0} = \frac{r_0}{\sigma^2} \qquad (26)$$

where δ_m represents the Kronecker delta.

Similarly, it is also possible to find the backward recursive equations. Due to the lack of space, we report only the final expressions here

$$q_{b,k-1}^{(m)} \propto \left[\sum_{i=0}^{M-1} P\left(a_k = e^{j\frac{2\pi}{M}i}\right) q_{b,k}^{(m+i)_{\mathrm{mod}\,M}}\right] \cdot e^{\left|z'_{b,k}e^{j\frac{2\pi}{M}m} + \frac{r_{k-1}}{\sigma^2}\right|}, \quad m = 0, \ldots, M-1 \qquad (27)$$

$$z_{b,k-1} = z'_{b,k} + \frac{r_{k-1}}{\sigma^2}\sum_m q_{b,k-1}^{(m)} e^{-j\frac{2\pi}{M}m} \qquad (28)$$

having defined $z'_{b,k} = z_{b,k}/(1+\sigma_\Delta^2|z_{b,k}|)$. The initial values of the backward coefficients are

$$q_{f,K}^{(m)} = \delta_m \qquad z_{f,K} = \frac{r_K}{\sigma^2}. \qquad (29)$$

Finally, substituting (16) and (17) into (15) and discarding irrelevant constants, the extrinsic information is evaluated as

$$\frac{P(a_k = e^{j\frac{2\pi}{M}i}|\mathbf{r})}{P(a_k = e^{j\frac{2\pi}{M}i})} = \sum_m \sum_\ell q_{f,k-1}^{(m)} q_{b,k}^{(\ell)} \cdot \mathrm{I}_0\left(\left|z'_{f,k-1} + z_{b,k} e^{j\frac{2\pi}{M}(\ell-m-i)}\right|\right). \quad (30)$$

In summary, this detection algorithm is based on three steps: a forward recursion in which, for each time epoch k, M real and one complex coefficients are evaluated based on the summations (24) and (25) of M terms, a backward recursion, based on (27) and (28), which proceeds similarly, and finally, a completion (30), which computes M values of extrinsic information, each through a sum of M^2 terms (although only M of them are numerically significant). This algorithm, which will be denoted as "algorithm based on Tikhonov parameterization" (*Tikh algorithm*), entails a limited complexity increase with respect to the known-phase MAP symbol detector [22]. In fact, to implement this latter algorithm, we must compute, in each recursion, M real quantities (instead of M real and 1 complex parameters, as in the *Tikh algorithm*) through summations of M terms and, in the completion, M extrinsic information values each through a sum of M terms.

V. Extension to RI Codes

Differential encoding belongs to the wider class of RI codes, introduced by Wei [11], [12], whose ideas were successively extended by other authors [13], [14]. For example, the necessary and sufficient conditions for a code and an encoder to posses such a property were presented in [13]. Although most of the RI concatenated codes in the literature have an inner differential component code [16]–[21], we now introduce the main concepts behind RI codes and encoders, as well as possible extensions of the proposed algorithms to RI codes.

We say that a sequence of modulation symbols is the rotated version of another sequence if every symbols of the first sequence is the rotated version of the symbol of the second sequence in the same position (i.e., componentwise rotation). A *code* is said to be RI if there exists an angle Θ (the *base angle*), such that every code sequence, rotated by any multiple of the base angle, is still a code sequence. Furthermore, if all the rotated versions of a given code sequence are associated to the same information sequence, the *encoder* is said to be RI. As demonstrated in [13], an RI encoder cannot be feedforward. This is a favorable property for serially concatenated schemes whose inner encoder must be recursive to have an interleaver gain.

Let us consider an S-state RI encoder whose input symbol, output symbol, and trellis state at discrete time k will be denoted by a_k, c_k, and s_k, respectively. The "next-state" and output functions of the encoder will be denoted by $s_{k+1} = \eta(s_k, a_k)$ and $c_k = \psi(s_k, a_k)$, respectively. Since the encoder is recursive, for any given state s_{k+1} and input symbol a_k, it is possible to obtain the state s_k, i.e., $s_k = \eta^{-1}(s_{k+1}, a_k)$. We will denote by $a^{(n)}$ the nth symbol of the input alphabet (with cardinality N) and by $c^{(m)}$ the mth symbol of the output alphabet (with cardinality $M \geq N$).

All trellis states of an RI encoder can be partitioned into orbits [13], [14] of size $P \geq 1$, where the orbit is defined as the set of "equivalent" states (in the sense specified in [13] and [14]) and thus, the total number of states of the encoder must be a multiple of P. It is worth noticing that, for the M-PSK differential encoder (which is clearly an RI encoder), $P = M$, that is, there exists only one orbit. Hence, the state space can be partitioned as $(\Sigma^{(0)}, \Sigma^{(1)}, \ldots, \Sigma^{(S/P-1)})$ where each $\Sigma^{(p)}$ collects all states belonging to the pth orbit. We will denote by $\sigma_0^{(p)}$ a reference state of the pth orbit (which can be whatever state).

Let us define $\rho_\ell(x) = xe^{j\ell\Theta}$, that is, $\rho_\ell(x)$ denotes a rotation of x by a multiple of the base angle. It is worth noticing that, from the definition of P, $\rho_P(\cdot)$ is an identity function. From the rotationally invariance of the encoder, with a proper labeling of the trellis states, it follows that [13]

$$\begin{aligned} \rho_\ell(s_{k+1}) &= \eta\left(\rho_\ell(s_k), a_k\right) \\ \rho_\ell(s_k) &= \eta^{-1}\left(\rho_\ell(s_{k+1}), a_k\right) \\ \rho_\ell(c_k) &= \psi\left(\rho_\ell(s_k), a_k\right) \end{aligned} \quad (31)$$

for any valid s_k, s_{k+1}, a_k and c_k.

Although the mathematical details are not reported here, we can prove that properties similar to the ones in Section III hold for a generic RI encoder. For example, in the forward recursion, for a given time epoch k, *a single message for each orbit* has to be updated. This message will be

$$\alpha_k^{(p)}(\theta_k) \triangleq p(\theta_k | s_{k+1} = \sigma_0^{(p)}, \mathbf{r}_0^k), \quad p = 0, 1, \ldots, \frac{S}{P} - 1. \quad (32)$$

The forward recursion turns out to be

$$\alpha_k^{(p)}(\theta_k) = \sum_{n=0}^{N-1} P\left(a_k = a^{(n)}\right) p\left(r_k | c^{(\overline{m})}, \theta_k\right) \cdot \int \alpha_{k-1}^{(\overline{q})}\left(\theta_{k-1} + \Theta\overline{\ell}\right) p(\theta_k|\theta_{k-1}) d\theta_{k-1} \quad (33)$$

where $\overline{\ell}$, and $\overline{q}$ are such that $\rho_{\overline{\ell}}(\sigma_0^{(\overline{q})}) = \eta^{-1}\left(\sigma_0^{(p)}, a^{(n)}\right)$ and $c^{(\overline{m})}$ is the modulation symbol associated with this transition, i.e., $c^{(\overline{m})} = \psi(\rho_{\overline{\ell}}(\sigma_0^{(\overline{q})}), a^{(n)})$. For the backward recursion and the completion, it turns out that

$$\beta_{k-1}^{(p)}(\theta_{k-1}) = \sum_{n=0}^{N-1} P\left(a_k = a^{(n)}\right) \int \beta_k^{(\overline{q})}(\theta_k + \Theta\overline{\ell}) \cdot p\left(r_k | c^{(\overline{m})}, \theta_k\right) p(\theta_k|\theta_{k-1}) d\theta_k \quad (34)$$

$$\frac{P\left(a_k = a^{(n)}|\mathbf{r}\right)}{P\left(a_k = a^{(n)}\right)} = \sum_{p=0}^{S/P-1} \iint \alpha_{k-1}^{(p)}(\theta_{k-1}) \beta_k^{(\overline{q})}(\theta_k + \Theta\overline{\ell}) \cdot p\left(r_k|c^{(\overline{m})}, \theta_k\right) p(\theta_k|\theta_{k-1}) d\theta_{k-1} d\theta_k \quad (35)$$

where, in this case, $\overline{\ell}$, and $\overline{q}$ are such that $\rho_{\overline{\ell}}(\sigma_0^{(\overline{q})}) = \eta(\sigma_0^{(p)}, a^{(n)})$ and $c^{(\overline{m})}$ is the modulation symbol associated with this transition, i.e., $c^{(\overline{m})} = \psi(\sigma_0^{(p)}, a^{(n)})$. It is worth noticing

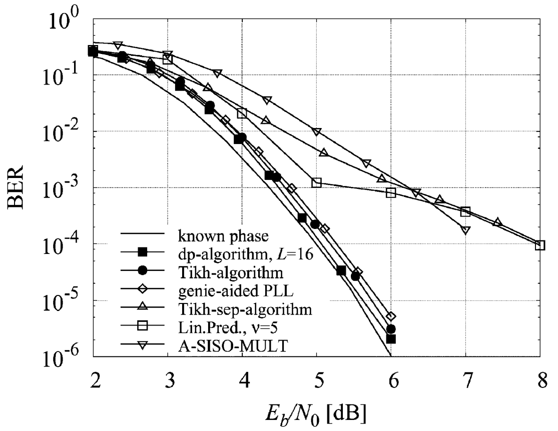

Fig. 2. Performance for a serially concatenated scheme composed by a rate-$2/3$ convolutional encoder and a differential encoder. The binary phase-shift keying (BPSK) modulation is considered and the phase noise follows the Wiener model with $\sigma_\Delta = 6°$.

that (33)–(35) represent a generalization to the case of general RI encoders of (13)–(15), respectively.

In this way, the *dp-algorithm* introduced in Section IV-A can be generalized to work with every RI encoder, irrespective of the used modulation format. On the other hand, the *Tikh algorithm* introduced in Section IV-B can be straightforwardly generalized as well, at least for single-orbit encoders (namely, when $S = P$). For RI encoders with more than one orbit, approximation (21) must be extended in order to derive a *Tikh-algorithm* for this scenario also (this approach is not pursued here for a lack of space).

VI. Numerical Results

In this section, the performance of the proposed algorithms is assessed by computer simulations in terms of bit error rate (BER) versus E_b/N_0, E_b being the received signal energy per information bit and N_0 the one-sided noise power spectral density. The sequence $\mathbf{a}$ is now assumed as a codeword of an outer channel code. After interleaving, these code symbols are then further encoded by means of an inner RI code.

In particular, in Fig. 2, we consider a serially concatenated scheme composed of a convolutional code (CC), an interleaver, and a differentially encoded BPSK. The CC is a rate-1/2 nonrecursive nonsystematic code with four states and generators $(5, 7)$ (octal notation). A uniform puncturing of its parity bits is used to obtain a rate-$2/3$ code. The codewords are composed of 16 200 bits. The proposed schemes are employed to perform joint detection and decoding of the differential code. The produced soft outputs are then exchanged, in an iterative way, with the soft-output decoder for the CC, and a maximum of 15 iterations of the overall scheme is allowed. The phase noise affecting the channel is modeled as a Wiener process with $\sigma_\Delta = 6°$. From Fig. 2, it can be observed that, despite the presence of this strong phase noise, the low-complexity *Tikh algorithm* exhibits only a negligible performance loss with respect to the known-phase case and to the practically optimal *dp-algorithm* (with $L = 16$).

For comparison purposes, we show (curve labeled *Tikh-sep-algorithm*) the performance of the algorithm in [10]. This algorithm is also based on Tikhonov parameterization, but performs a phase tracking separate from the decoding of the differential code, and thus does not exploit the differential code constraints. In order for this algorithm to bootstrap, a pilot symbol for every 20 code symbols has been inserted in the frame, thus decreasing the effective information rate. This results in an increase in the required SNR of about 0.21 dB. Despite the presence of pilot symbols that help the algorithm in tracking the phase changes, the relevant performance loss is significant. The reason is related to the weakness of the concatenated code. In fact, a stronger code would be able to provide, after a few iterations, reliable decisions on some code symbols, thus providing additional pilot symbols to help the phase tracking.

We also show the performance of two other algorithms previously proposed in the literature. The first one is based on linear prediction [4] and is an improvement of the noncoherent algorithm described in [3]. It works on an expanded trellis of $M^{\nu-1}$ states ($2^{\nu-1}$ in this case of BPSK), where ν is the prediction order. Although complexity reduction techniques such as those described in [27] can be adopted, the full-complexity receiver has been considered, since we are interested in the best achievable performance. In Fig. 2, the case of $\nu = 5$ is considered, and practically, no performance improvement is obtained with larger values of ν. As can be observed, although this algorithm still exploits the knowledge of the channel statistics and works on an expanded trellis, its performance is far from that of the proposed algorithms. A large performance loss is also observed for another algorithm in the literature, namely, the *A-SISO-MULT* algorithm proposed in [5]. The performance curve shown in Fig. 2 refers to the *A-SISO-MULT* algorithm working on the trellis of the differential code. Perfect knowledge of the channel phase at the beginning and end of each codeword is assumed. This can be obtained, at least in the absence of phase noise, by inserting a sufficient number of known symbols although, for simplicity, the relevant decrease in the effective information rate has not been considered. This algorithm does not employ the knowledge of the channel phase statistics. However, the performance loss is not due to this fact, but due to the error propagation in the multiple phase estimates based on the per-survivor processing. In fact, if we assume, in the phase estimation process, a perfect knowledge of the transmitted symbols, that is, we consider a perfectly initialized *genie-aided* phase-locked loop (PLL) with an optimized equivalent bandwidth, we practically obtain the same performance obtained with the proposed algorithms whose excellence is, hence, further demonstrated.

In Fig. 3, we consider a nonsystematic rate-1/2 irregular RA code [20], [21] mapped on a quartenary PSK (QPSK) modulation before the differential encoder. The codewords have length of 880 symbols, and a maximum of 25 iterations is allowed. The RA code has been designed following the approach in [28], and the relevant degree distributions are reported in Table I. The statistics of the phase noise are those characterizing consumer-grade equipment operating at the carrier frequency of the second-generation digital video broadcasting satellite system (DVB-S2). A complete description of this DVB-S2 compliant phase noise model can be found in [10] and references therein. In Fig. 3, the case of a baud rate of 10 MBd has been considered. Although the Wiener model does not apply to this case, the proposed algorithms work well with a

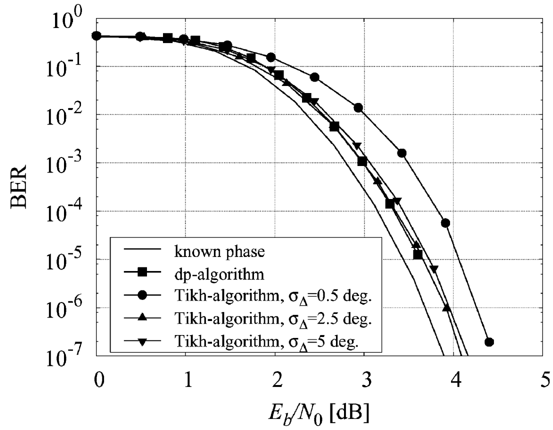

Fig. 3. Performance for an optimized RA code mapped over a QPSK modulation. The DVB-S2 phase noise model is considered (10 MBd).

TABLE I
DEGREE DISTRIBUTIONS OF THE EMPLOYED RA CODE

$d_{v,1} = 3$	$a_{v,1} = 0.8126$	$d_{c,1} = 1$	$a_{c,1} = 0.2000$
$d_{v,2} = 4$	$a_{v,2} = 0.1375$	$d_{c,2} = 3$	$a_{c,2} = 0.8000$
$d_{v,3} = 15$	$a_{v,3} = 0.0294$		
$d_{v,4} = 20$	$a_{v,4} = 0.0204$		

properly optimized value of $\sigma_\Delta = 2.5°$. From the figure, it can be observed that the *Tikh-algorithm* has about the same performance of the much more complex *dp-algorithm* (with $L = 32$), and the two algorithms exhibit a loss with respect to the known phase case of only 0.2 dB. We would like to point out that this very good result has been obtained without the insertion of pilot symbols. The other algorithms in the literature are not reported here, since they are characterized by a larger performance loss.

We would like to discuss the role of the parameter σ_Δ. In the derivation of the proposed algorithms this parameter has been assumed to be known to the receiver, whereas in a practical implementation, only its nominal value is known. However, the phase noise statistics depend on the characteristics of the local oscillators whose variations are limited and, on the other hand, the sensitiveness of the receiver to this parameter is very low. This is shown for the *Tikh algorithm* with reference to the system whose performance is reported in Fig. 3. In fact, as already mentioned, the optimal value of σ_Δ, optimized by means of computer simulations, is $\sigma_\Delta = 2.5°$. However, the performance loss for $\sigma_\Delta = 5°$ is negligible, whereas it is lower than 0.3 dB for $\sigma_\Delta = 0.5$ degrees (see Fig. 3). As a final remark, we would like to remember that, in other receivers in the literature, there are parameters to be optimized for the channel at hand. For example, the equivalent bandwidth in a classical PLL must be optimized for the phase noise under consideration.

We now consider another RI serially concatenated scheme recently proposed in [17] and [18], which is again based on an inner differential encoder. This scheme is, in brief, composed of an outer rate-$2/3$ parity code, an "*S*-random" interleaver with parameter $S = 10$ [29] and, after a mapping onto an 8-PSK constellation, a differential encoder. The interleaver size is of 15 000 bits and the employed 8-PSK constellation exploits a labeling properly optimized in order to reduce the error floor [17], [18].

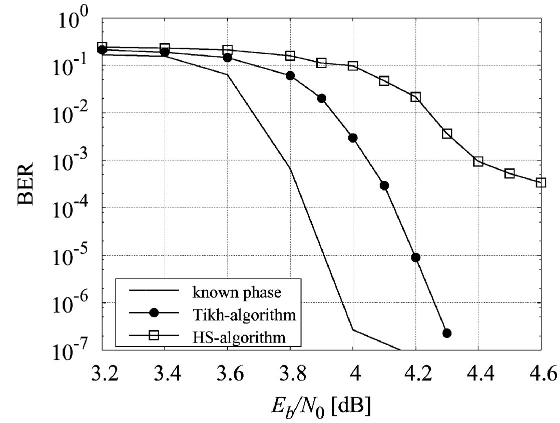

Fig. 4. Performance for a serially concatenated scheme composed by a rate-$2/3$ parity code mapped onto an 8-PSK constellation and a differential encoder. The DVB-S2 phase noise model is considered (10 MBd).

Up to 50 iterations are allowed for the iterative decoder. The phase noise model is identical to the one considered in Fig. 3, namely, the DVB-S2 compliant phase noise model assuming a baud rate of 10 MBd. A simple algorithm for detection in the presence of phase uncertainties has also been proposed in [17], [18]. This algorithm will be denoted as "Howard–Schlegel algorithm" (*HS-algorithm*) and is an improvement, for time-varying channels, of the algorithm based on expectation–maximization in [8]. Although this algorithm performs a detection separate from the decoding of the differential code, in [17] and [18], it is shown that no pilot symbols are necessary to bootstrap. For the considered phase noise, a decay factor [17], [18] $\alpha = 0.99$, optimized by simulations, has been used. In Fig. 4, we show the performance of the proposed *Tikh-algorithm* (with $\sigma_\Delta = 0.6°$) and that of the *HS-algorithm*. The *Tikh-algorithm* clearly outperforms the one by Howard and Schlegel in this scenario, since this latter algorithm exhibits a very large error floor and is not able to get down below a BER of 10^{-4}.

To further highlight the properties of the compared detection schemes, we also use the recently proposed technique known as extrinsic information transfer (EXIT) chart analysis [30], which is a powerful tool designed to evaluate the system performance of concatenated schemes in the waterfall region. This technique has been widely used to numerically evaluate (in a way much faster than the Monte Carlo BER simulations) the convergence threshold of iteratively decoded concatenated systems [10] and to design error-correcting codes [28], [31] with low thresholds. Fig. 5 represents the average mutual information (AMI) at the output of the inner decoder ($I_{\mathrm{out,inner}}$) as a function of the AMI at its input ($I_{\mathrm{in,inner}}$), for the case of perfect knowledge of the channel, the *Tikh-algorithm*, and the *HS-algorithm*. All curves except one (specified) refers to $E_b/N_0 = 4$ dB. The EXIT chart is completed by the curve of the input AMI of the decoder for the outer code ($I_{\mathrm{in,outer}}$) as a function of the AMI at its output ($I_{\mathrm{out,outer}}$). The EXIT analysis clearly shows that, at $E_b/N_0 = 4$ dB, both the known-phase detector and the proposed *Tikh-algorithm* are above threshold (i.e., open tunnel), while the tunnel is closed for the algorithm proposed in [17], [18] also at $E_b/N_0 = 4.5$ dB. It is also interesting to note that, for large enough extrinsic information coming from the outer decoder,

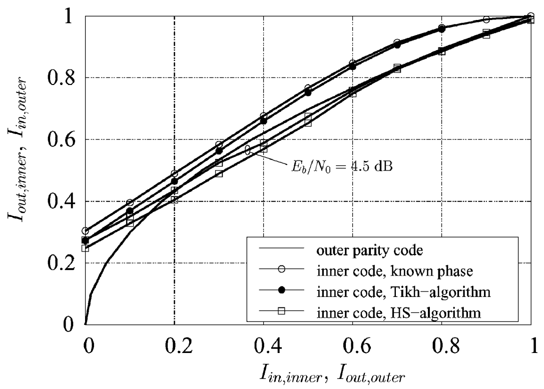

Fig. 5. EXIT chart for the same serially concatenated scheme of Fig. 4 and the same phase noise model. All curves except one (specified) were obtained at $E_b/N_0 = 4$ dB.

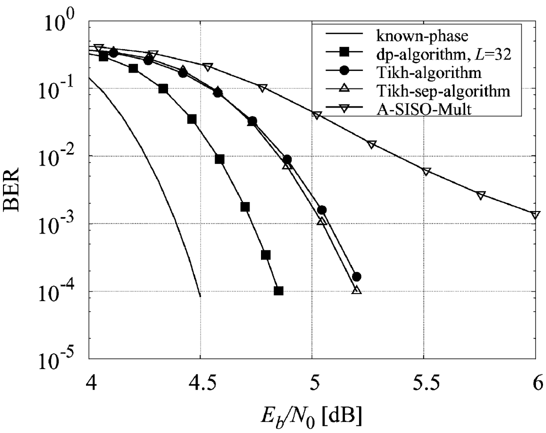

Fig. 6. Performance for a SC-TCM scheme composed of a rate-$2/3$ CC and an RI 16-QAM TCM encoder. The phase noise follows the Wiener model with $\sigma_\Delta = 1°$.

when all symbols practically become pilots, the proposed *Tikh-algorithm* behaves approximately as the known-phase case, that is, the phase estimation becomes extremely reliable. On the other hand, this valuable property is missing in the *HS-algorithm*.

We now consider the extension of the proposed algorithms to a generic RI code. In Fig. 6, we consider the RI serially concatenated TCM denoted as "Code 2" in [14, p. 2004]. In this scheme, a rate-$2/3$ outer convolutional code is concatenated, through an "S-random" interleaver with parameter $S = 10$ [29], to an RI-TCM code over the 16-quadrature amplitude modulation (16-QAM). The overall system efficiency is 2 bits per channel use, and the length of the employed interleaver is of 12 288 bits. At the receiver, a maximum of 25 iterations is allowed. In the figure, for a Wiener phase noise with $\sigma_\Delta = 1°$, we show the performance of the *dp-algorithm* using $L = 32$, and also of the *Tikh-algorithm* and the *Tikh-sep-algorithm* when one pilot for every 20 code symbols has been inserted to allow the algorithm bootstrap, and of the A-SISO-MULT working on the code trellis and with a perfect initialization of the forward and backward recursions. Due to the approximations involved in the derivation of the proposed algorithms, at the receiver side, it is better to adopt $\sigma_\Delta = 2.2°$. As can be seen, our algorithms outperform the A-SISO-MULT that has a high error floor, and despite the absence of the pilot symbols, for this powerful code, the *Tikh-algorithm* has a much better performance than that of the *Tikh-sep-algorithm*. The *Tikh-algorithm* exhibits a loss of only 0.05 dB with respect to the *dp-algorithm* with a significant complexity reduction, since only one complex and four real parameters must be updated in the forward and backward recursions.

VII. Conclusion

In this paper, the problem of MAP symbol detection for RI trellis-coded modulations transmitted over channels affected by phase noise has been faced. A simplified, although exact, version of the algorithm has been derived based on a reduced number of forward–backward estimators of the phase pdf and a final completion. For the practical implementation of the forward–backward estimators, two algorithms have been proposed. The first one is based on the phase discretization and becomes optimal for a large enough number of discretization levels. To reduce the computational complexity, some approximations have been introduced in order to derive a new algorithm which exhibits a very good performance and a lower complexity. For serially concatenated RI schemes, an accurate performance comparison with previously proposed solutions as well as with ideal receivers has been accomplished.

Appendix

In this appendix, we prove the three properties introduced in Section III. For the first two properties, we concentrate on the forward recursion, since the extension to the backward recursion is trivial. The forward recursion can be expressed as [by using (11) in (8)]

$$\begin{aligned} P(c_k|\mathbf{r}_0^k)p(\theta_k|c_k,\mathbf{r}_0^k) &= p(r_k|c_k,\theta_k)\sum_{a_k}P(a_k)\int p(\theta_k|\theta_{k-1}) \\ \cdot P(c_{k-1} = c_k a_k^*|\mathbf{r}_0^{k-1})p(\theta_{k-1}|c_{k-1} &= c_k a_k^*, \mathbf{r}_0^{k-1})d\theta_{k-1}. \end{aligned} \quad \text{(A1)}$$

The first two properties can be demonstrated by induction. First of all, $P(c_0|r_0) = 1/M$, since the initial state of the differential encoder is supposed to be unknown to the receiver, and $p(\theta_0|c_0, r_0) \propto p(r_0|c_0,\theta_0)$. Now, supposing that the two properties hold at time $k-1$, we can easily prove that they also hold at time k by evaluating (A1) for $\bar{c}_k = c_k e^{j(2\pi/M)i}$, exploiting Property 1, for $P(c_{k-1}|\mathbf{r}_0^{k-1})$, and Property 2, for $p(\theta_{k-1}|c_{k-1},\mathbf{r}_0^{k-1})$, by applying a suitable change of variable, and noting that, for each angle ϵ, $p(\theta_k|\theta_{k-1}-\epsilon) = p(\theta_k+\epsilon|\theta_{k-1})$.

We now consider the third property, involving the completion (10). By exploiting the first property, the extrinsic information can be written as the sum of M terms of the form

$$\iint p(\theta_{k-1}|c_{k-1},\mathbf{r}_0^{k-1})p(\theta_k|c_k = c_{k-1}a_k,\mathbf{r}_k^K)\cdot p(\theta_k|\theta_{k-1})\, d\theta_k\, d\theta_{k-1}. \quad \text{(A2)}$$

Let us now consider the term corresponding to the value $\bar{c}_{k-1} = c_{k-1}e^{j(2\pi/M)i}$. Since a_k is fixed, and by exploiting the second property, this term is clearly equal (after a change of variables in the integrals) to the term (A2) for any value of i.

ACKNOWLEDGMENT

The authors would like to thank Prof. C. Schlegel and Dr. S. Howard for some useful discussions about the *HS-algorithm*.

REFERENCES

[1] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1376–1385, Sep. 1999.

[2] M. Peleg, S. Shamai (Shitz), and S. Galán, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel," *Proc. IEE Commun.*, vol. 147, pp. 87–95, Apr. 2000.

[3] G. Colavolpe, G. Ferrari, and R. Raheli, "Noncoherent iterative (turbo) decoding," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1488–1498, Sep. 2000.

[4] G. Ferrari, G. Colavolpe, and R. Raheli, "On linear predictive detection for communications with phase noise and frequency offset," *IEEE Trans. Veh. Techol.*, vol. 56, no. 4, pp. 2073–2085, Jul. 2007.

[5] A. Anastasopoulos 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.

[6] J. Dauwels and H.-A. Loeliger, "Joint decoding and phase estimation: An exercise in factor graphs," in *Proc. IEEE Int. Symp. Inf. Theory*, Yokohama, Japan, Jun./Jul. 2003, p. 231.

[7] R. Nuriyev and A. Anastasopoulos, "Pilot-symbol-assisted coded transmission over the block-noncoherent AWGN channel," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 953–963, Jun. 2003.

[8] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, Apr. 2004.

[9] G. Colavolpe, "On LDPC codes over channels with memory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1757–1766, Jul. 2006.

[10] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, Sep. 2005.

[11] L.-F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—Part I: 180°," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 5, pp. 659–671, Sep. 1984.

[12] L.-F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—Part II: Nonlinear codes," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 5, pp. 672–686, Sep. 1984.

[13] M. D. Trott, S. Benedetto, R. Garello, and M. Mondin, "Rotational invariance of trellis codes—Part I: Encoders and precoders," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 751–765, May 1996.

[14] R. Nuriyev and A. Anastasopoulos, "Rotationally invariant and rotationally robust codes for the AWGN and the noncoherent channel," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2001–2010, Dec. 2003.

[15] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[16] W. Liu and S. Wilson, "Rotationally invariant concatenated (turbo) TCM codes," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 1999, vol. 1, pp. 32–36.

[17] S. Howard and C. Schlegel, "Differentially-encoded turbo coded modulation with APP channel estimation," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, CA, Nov. 2003, pp. 1761–1765.

[18] S. L. Howard and C. Schlegel, "Differential turbo coded modulation with APP channel estimation," *IEEE Trans. Commun.,* vol. 54, no. 8, pp. 1397–1406, Aug. 2007.

[19] M. Franceschini, G. Ferrari, R. Raheli, and A. Curtoni, "Serial concatenation of LDPC codes and differential modulations," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1758–1768, Sep. 2005.

[20] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. Annu. Allerton Conf. Commun., Control Comput.*, Urbana, IL, Sep.1998, pp. 201–210.

[21] H. Jin, A. Khandekar, and R. J. McEliece, "Irregular repeat-accumulate codes," presented at the Int. Symp. Turbo Codes Relat. Topics, Brest, France, Sep. 2000.

[22] 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. 20, no. 2, pp. 284–287, Mar. 1974.

[23] A. P. Worthen and W. E. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.

[24] 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.

[25] R.-R. Chen, 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.

[26] A. Barbieri, G. Colavolpe, and G. Caire, "Joint iterative detection and decoding in the presence of phase noise and frequency offset," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 171–179, Jan. 2007.

[27] 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.

[28] S. ten Brink and G. Kramer, "Design of repeat–accumulate codes for iterative detection and decoding," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2764–2772, Nov. 2003.

[29] D. Divsalar and F. Pollara, "Turbo codes for PCS applications," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, 1995, pp. 54–59.

[30] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[31] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, Apr. 2004.

Alan Barbieri was born in Parma, Italy, in 1979. He received the Dr.Ing. degree (*cum laude*) in telecommunications engineering and the Ph.D. degree in information technology from the University of Parma, Parma, Italy, in 2003 and 2007, respectively.

He is currently a Research Associate in the Dipartimento di Ingegneria dell'Informazione (DII), University of Parma. Since March 2007, he has also been with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, as a Postdoctoral Visiting Scholar. His current research interests include digital transmission theory and information theory, with particular emphasis on channel coding, iterative joint detection and decoding algorithms, estimation of unknown parameters, and algorithms for synchronization. He had participated in several research projects funded by the European Space Agency (ESA-ESTEC) and other important telecommunications companies.

Dr. Barbieri was the recipient of the Premio Conti for the year 2003, as the Best Graduate in information engineering at the University of Parma, 2003.

Giulio Colavolpe (S'96–M'00) was born in Cosenza, Italy, in 1969. He received the Dr. Ing. degree (*cum laude*) in telecommunications engineering 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. In 2000, he was a Visiting Scientist at the Institut Eurécom, Valbonne, France. His current research interests include digital transmission theory, adaptive signal processing, channel coding, and information theory. His research activity has led to numerous 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, with Applications to Wired and Storage Systems* (New York: Wiley, 2004).

Dr. Colavolpe received the Best Paper Award at the 13th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2005). He is also the Principal Investigator of several research projects funded by the European Space Agency (ESA-ESTEC) and other important telecommunications companies.