Algorithms for Iterative Decoding in the Presence of Strong Phase Noise

Giulio Colavolpe, Alan Barbieri, *Student Member, IEEE*, and Giuseppe Caire, *Fellow, IEEE*

***Abstract*—We present two new iterative decoding algorithms for channels affected by strong phase noise and compare them with the best existing algorithms proposed in the literature. The proposed algorithms are obtained as an application of the sum-product algorithm to the factor graph representing the joint *a posteriori* probability mass function of the information bits given the channel output. In order to overcome the problems due to the presence in the factor graph of continuous random variables, we apply the method of *canonical distributions*. Several choices of canonical distributions have been considered in the literature. Well-known approaches consist of discretizing continuous variables or treating them as jointly Gaussian, thus obtaining a Kalman estimator. Our first new approach, based on the Fourier series expansion of the phase probability density function, yields better complexity/performance tradeoff with respect to the usual discretized-phase method. Our second new approach, based on the Tikhonov canonical distribution, yields near-optimal performance at very low complexity and is shown to be much more robust than the Kalman method to the placement of pilot symbols in the coded frame. We present numerical results for binary LDPC codes and LDPC-coded modulation, with particular reference to some phase-noise models and coded-modulation formats standardized in the next-generation satellite Digital Video Broadcasting (DVB-S2). These results show that our algorithms achieve near-coherent performance at very low complexity without requiring any change to the existing DVB-S2 standard.**

***Index Terms*—Channels with memory, factor graphs (FGs), iterative detection/decoding, low-density parity-check (LDPC) codes, phase-noise, sum-product algorithm (SPA), Tikhonov parameterization.**

I. Introduction

THE FACTOR GRAPH (FG) representation and the sum-product algorithm (SPA) provide a general and powerful framework to derive low-complexity Bayesian detection and decoding algorithms [1]. In this paper, we make use of this framework to derive efficient algorithms for iterative decoding in additive white Gaussian noise (AWGN) channels affected by phase noise. We construct the FG corresponding to the joint *a posteriori* probability distribution of the information message bits and of the random channel parameter (phase noise in our case), given the received signal. Then, we let the SPA compute the posterior marginal distributions of the information bits. Bit-by-bit decisions are then made, based on the resulting posterior marginals. The FG takes into account the probabilistic structure of the channel parameters, such that expectation over the unknown parameters is implicitly performed by the SPA as part of the marginalization. The posterior marginal probabilities computed by the SPA are *exact* if the underlying FG is cycle-free. In this case, the bit-by-bit decision is optimal, i.e., it minimizes the average bit-error probability. More often, the underlying FG has cycles and the resulting SPA is inherently iterative. In this case, the SPA does not yield in general the optimal MAP decision rule. Nevertheless, the iterative SPA has proven to provide very good performance in several problems and, therefore, it can be regarded as a viable low-complexity solution when the optimal decision rule is just too complex to be implemented in practice. Since the resulting algorithms are naturally iterative, they are particularly suited to the decoding of codes such as low-density parity-check (LDPC) and turbo codes, whose decoding algorithms are typically iterative (and suboptimal) even in the fully coherent setting (all channel parameters known).

Iterative decoding algorithms for channels with unknown phase have attracted an increasing interest in the recent literature. The algorithms developed in [2]–[7] are designed for noncoherent decoding of turbo codes and can be applied to LDPC codes only if trellis-based separate detection is performed. In particular, in [2] and [7], receivers for both the block-constant phase model and a discretized random-walk phase model are developed by using a phase discretization approach. In [8], the use of FGs that include both the code constraints and the channel parameter statistics is advocated in a very general setting. By specializing the approach of [8] to particular channel phase statistics, several algorithms for noncoherent detection/decoding have been proposed. In [9] and [10], LDPC ensemble optimization via density evolution is considered for a very simplified block-constant phase model quantized over the two levels 0 and π. In [11], a constant and a random-walk phase noise model with Gaussian increments are considered and approximations of the SPA are derived and evaluated. In [12], the messages in the SPA related to continuous random variables are replaced by Dirac impulses located at estimated values for the corresponding variable and different estimation methods are examined. In [13], a phase model where the unknown carrier phase is constant over a block of N symbols and independent from block to block is considered, the channel parameters are not explicitly introduced in the FG, and the power allocation to the pilot symbols

Manuscript received March 30, 2004; revised December 6, 2004 and February 25, 2005. This work was supported in part by the European Space Agency, ESA-ESTEC, Noordwijk, The Netherlands, under Contract 17337/NL/LvH. This paper was presented in part at the IEEE Symposium on Information Theory (ISIT'04), Chicago, IL, June–July 2004.

G. Colavolpe and A. Barbieri are with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, Parma 43100, Italy (e-mail: giulio@unipr.it; barbieri@tlc.unipr.it).

G. Caire is with the Mobile Communications Group, Institut Eurécom, Valbonne 6560, France (e-mail: giuseppe.caire@eurecom.fr).

Digital Object Identifier 10.1109/JSAC.2005.853813

is optimized by using density evolution. Finally, in [14], based on an assumption of memory truncation, a general Bayesian approach to LDPC decoding without the explicit representation of the channel parameters into the FG has been proposed and applied to different channel models.

A non-Bayesian approach is adopted in [15]–[21]. In [15]–[20], the concept of *soft-decision-directed* estimation is introduced. The channel parameters are estimated by using the expectation-maximization (EM) algorithm [15]–[19] or an ad hoc procedure [20] and the estimation algorithm is embedded into the iterative decoding process. Generally speaking, non-Bayesian approaches consider the channel phase as a deterministic unknown constant. Tracking time-variations, such as in the case of phase noise, is obtained by using some heuristic sliding window adaptation. As a matter of fact, while the non-Bayesian schemes may be suited for the block-constant phase model, their performance degrades significantly in the presence of phase noise since the algorithms are not designed by exploiting the statistical knowledge of the phase time-variations.

In this paper, we follow the FG/SPA framework of [8], and we focus on the random-walk phase noise model with Gaussian increments, as in [11]. While the SPA is well-suited to handle discrete random variables, characterized by a probability mass function (pmf), the channel parameters are typically continuous random variables, characterized by a probability density function (pdf). The SPA for continuous random variables involves integration and computation of continuous pdfs, and it is not suited for direct implementation. A solution for this problem is suggested in [8] and consists of the use of *canonical distributions*, i.e., the pdfs computed by the SPA are constrained to be in a certain "canonical" family, characterized by some parameterization. Hence, the SPA reduces to propagating and updating the parameters of the pdf rather than the pdf itself. Beyond this general idea, several different algorithms can be obtained depending of the choice of the canonical distribution family. These approximations of the SPA, albeit all derived from the same standard approach, offer different complexity and performance. Therefore, finding good canonical distribution parameterizations suited to the problem at hand is the key step in algorithm design.

The most straightforward parameterization is based on the discretization of the parameter space [2], [7], [11]. In the case of phase-noise, the application of the SPA computation rules to the discretized random-walk yields a BCJR algorithm that operates on a trellis representation of the phase noise trajectories. Another well-known approach is based on modeling the phasor process and the channel observations as jointly Gaussian, thus obtaining a modified version of the well-known Kalman smoother [11].

In this paper, we propose two new approaches based on Fourier and Tikhonov parameterizations, respectively. The Fourier approach explicitly exploits the fact that the phase pdf is periodic. Hence, for high-order modulations it yields better complexity/performance tradeoff than straightforward discretization. The Tikhonov approach yields a one-dimensional forward–backward recursion that can be regarded (roughly speaking) as a nonlinear version of the Kalman smoother. Remarkably, its performance is nearly as good as the discretized phase approach (nearly optimal) with considerable less complexity, and it is much more robust than the Kalman smoother to the placement of pilot symbols.[1] As a matter of fact, the newly proposed Tikhonov parameterization yields an algorithm with unprecedented performance/complexity tradeoff, thus setting a new state-of-the art in joint phase synchronization and decoding.

The reminder of this paper is organized as follows. Section II is mostly tutorial. It introduces the channel model and presents the derivation of the exact SPA. In Section III, we briefly review the discretized-phase BCJR and the Kalman methods, that are used as terms of comparison of the new proposed algorithms. Section IV presents the details of the new algorithms. A complexity comparison is carried out in Section V. Finally, in Section VI we present some numerical results and in Section VII, we point out some concluding remarks.

II. System Model and Exact Sum-Product Algorithm (SPA)

We consider the transmission of a sequence of complex modulation symbols $\mathbf{c} = (c_0, c_1, \ldots, c_{K-1})$ over an AWGN channel affected by carrier phase noise. Symbols c_k are linearly modulated. Assuming Nyquist transmitted pulses, matched filtering, and phase variations slow enough so as no intersymbol interference arises, the discrete-time baseband complex equivalent channel model at the receiver is given by

$$r_k = c_k e^{j\theta_k} + n_k, \quad k = 0, \ldots, K-1. \tag{1}$$

We assume that the sequence $\mathbf{c}$ is a codeword of the channel code $\mathcal{C}$ constructed over an M-ary modulation constellation $\mathcal{X} \subset \mathbb{C}$. We include possible pilot symbols (known to the receiver) and/or possible differential encoding as a part of the code $\mathcal{C}$. The vector of noise samples $\mathbf{n} = (n_0, n_1, \ldots, n_{K-1})$ has independent identically distributed (i.i.d.), complex circularly symmetric components, with $n_k \sim \mathcal{N}_{\mathbb{C}}(0, 2\sigma^2)$.[2] The vector of channel phases $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_{K-1})$ is random, unknown to both transmitter and receiver, and statistically independent of $\mathbf{c}$ and $\mathbf{n}$.

A common model for the phase noise process $\{\theta_k\}$ is the random-walk (Wiener) model described by

$$\theta_k = \theta_{k-1} + \Delta_k \tag{2}$$

where $\{\Delta_k\}$ is a white real Gaussian process with $\Delta_k \sim \mathcal{N}(0, \sigma_\Delta^2)$. Under this assumption and assuming $\theta_0 \sim \mathrm{Uniform}[0, 2\pi)$, it follows that

$$p(\theta_k | \theta_{k-1}, \theta_{k-2}, \ldots, \theta_0) = p(\theta_k | \theta_{k-1}) = p_\Delta(\theta_k - \theta_{k-1}) \tag{3}$$

[1]We found that a minimum of pilot symbols to bootstrap the iterative decoder is necessary to all these algorithms in the case of strong phase noise and long block length. This will be illustrated briefly in Section VI from an extrinsic information transfer (EXIT) chart argument. We hasten to say that in the case of slowly varying phase and short block length these algorithm might work also without pilot symbols, in a fully noncoherent way.

[2]A complex circularly symmetric (respectively, real) Gaussian random vector $\mathbf{v}$ with mean $\mathbf{m}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathbf{v} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{m}, \boldsymbol{\Sigma})$ [respectively, by $\mathbf{v} \sim \mathcal{N}(\mathbf{m}, \boldsymbol{\Sigma})$]. We denote the multivariate complex circularly symmetric (respectively, real) Gaussian pdf with mean $\mathbf{m}$, covariance matrix $\boldsymbol{\Sigma}$ and argument $\mathbf{x}$ by $g_{\mathbb{C}}(\mathbf{m}, \boldsymbol{\Sigma}, \mathbf{x})$ [respectively, by $g(\mathbf{m}, \boldsymbol{\Sigma}, \mathbf{x})$].

where we define $p_\Delta(\varphi)$ as the pdf of the increment $\Delta_k \bmod 2\pi$, i.e.,

$$p_\Delta(\varphi) \triangleq \sum_{\ell=-\infty}^{\infty} g\left(0, \sigma_\Delta^2, \varphi - \ell 2\pi\right) \tag{4}$$

in any interval of length 2π. The Wiener phase noise model will be considered in the following as a working assumption in order to derive efficient iterative decoding algorithms. This assumption will be relaxed in Section VI, where we apply our algorithms to the DVB-S2-compliant European Space Agency (ESA) model described in [22] and [23] (see Section VI).

Without loss of generality, we assume that the code $\mathcal{C}$ admits an encoding function $\mu_\mathcal{C} : \mathbb{F}_2^B \to \mathcal{X}^K$, mapping binary information messages $\mathbf{b} \in \mathbb{F}_2^B$ into the codewords. The optimal decision rule that minimizes the average bit-error probability is given by

$$\widehat{b}_i = \arg\max_{b_i \in \mathbb{F}_2} P(b_i|\mathbf{r}) \tag{5}$$

where $P(b_i|\mathbf{r})$ denotes the *a posteriori* pmf for the ith information bit b_i given the received signal vector $\mathbf{r} = (r_0, \ldots, r_{K-1})$. Let $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{r})$ denote the joint posterior probability distribution function[3] of the information bits and of the phase noise vector $\boldsymbol{\theta}$ given $\mathbf{r}$. Clearly, the desired $P(b_i|\mathbf{r})$ can be obtained by marginalizing $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{r})$ with respect to $\boldsymbol{\theta}$ and to all b_j for $j \neq i$. This can be accomplished in an approximated but low-complexity way by the SPA applied on the FG of $P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{r})$, as illustrated in the following.

We assume that the reader is familiar with the FG/SPA framework (that can be found, for example, in the excellent tutorial paper [1]). Therefore, for the sake of space limitation, we will not recall here this background. From the definition of the encoding function $\mu_\mathcal{C}$ and the channel model (1), we obtain the factorization[4]

$$\begin{aligned} P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{r}) &\propto P(\mathbf{b}) p(\boldsymbol{\theta}) p(\mathbf{r}|\boldsymbol{\theta}, \mathbf{b}) \\ &\propto \chi\left[\mathbf{c} = \mu_\mathcal{C}(\mathbf{b})\right] p(\boldsymbol{\theta}) p\left(\mathbf{r}|\boldsymbol{\theta}, \mathbf{c} = \mu_\mathcal{C}(\mathbf{b})\right) \\ &\propto \chi\left[\mathbf{c} = \mu_\mathcal{C}(\mathbf{b})\right] p(\boldsymbol{\theta}) \prod_{k=0}^{K-1} p(r_k|c_k, \theta_k) \\ &\propto \chi\left[\mathbf{c} = \mu_\mathcal{C}(\mathbf{b})\right] p(\boldsymbol{\theta}) \prod_{k=0}^{K-1} f_k(c_k, \theta_k) \end{aligned} \tag{6}$$

where we have used the fact that the output signal pdf $p(\mathbf{r})$ does not depend on $\mathbf{b}$, that the information bits are uniform and i.i.d., therefore, $P(\mathbf{b}) = 2^{-B}$, that the AWGN channel for given $\boldsymbol{\theta}$ is memoryless. We have also defined the functions

$$\begin{aligned} f_k(c_k, \theta_k) &\triangleq \exp\left\{\frac{1}{\sigma^2}\mathrm{Re}\left[r_k c_k^* e^{-j\theta_k}\right] - \frac{|c_k|^2}{2\sigma^2}\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}|r_k - c_k e^{j\theta_k}|^2\right\} \end{aligned} \tag{7}$$

[3]We use the term probability distribution function to denote a continuous pdf with some discrete probability masses. For a probability distribution function, we still use the symbol $P(.)$.

[4]In this paper, we use extensively the proportionality relationship $f \propto g$, indicating that $f = ag$ for some real constant a, since the SPA is defined up to scaling its messages by positive factors, independent of the variables represented in the graph.

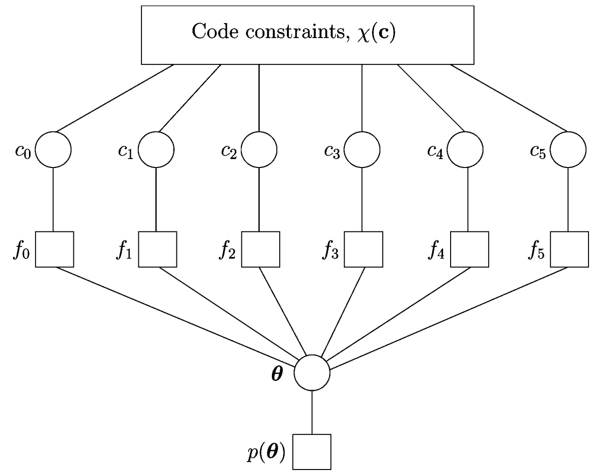

Fig. 1. Factor graph corresponding to (6) for $K = 6$.

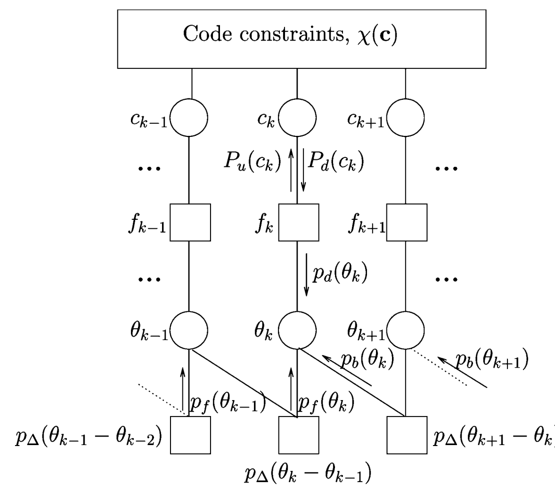

Fig. 2. Factor graph corresponding to (8).

and the code indicator function $\chi[\mathbf{c} = \mu_\mathcal{C}(\mathbf{b})]$, equal to 1 if $\mathbf{c}$ is the codeword corresponding to $\mathbf{b}$ and to zero, otherwise. The FG corresponding to (6) is shown in Fig. 1 for $K = 6$.

Under the assumption of first-order Markov model [see (3)] for the phase noise, we can further factor the term $p(\boldsymbol{\theta})$ as $p(\boldsymbol{\theta}) = p(\theta_0) \prod_{k=1}^{K-1} p_\Delta(\theta_k - \theta_{k-1})$ obtaining

$$\begin{aligned} P(\mathbf{b}, \boldsymbol{\theta}|\mathbf{r}) \propto \chi\left[\mathbf{c} = \mu_\mathcal{C}(\mathbf{b})\right] p(\theta_0) \prod_{k=1}^{K-1} p_\Delta(\theta_k - \theta_{k-1}) \\ \cdot \prod_{k=0}^{K-1} f_k(c_k, \theta_k). \end{aligned} \tag{8}$$

The corresponding FG is sketched in Fig. 2 and represents the starting point for the development of the proposed algorithms.

The SPA applied to the FG in the upper box, corresponding to the code constraints, consists of the well-known standard belief propagation whose efficient implementation depends on the structure of the code $\mathcal{C}$ and needs no details here. Hence, we shall concentrate on the SPA message propagation in the lower part of the graph. Omitting for simplicity of notation the explicit reference to the current iteration, let us denote by $P_d(c_k)$ the message from variable node c_k to factor node f_k, and by $P_u(c_k)$ the message in the opposite direction (see Fig. 2). The message $p_d(\theta_k)$ from factor node f_k to variable node θ_k can be expressed as

$$p_d(\theta_k) \propto \sum_{x\in\mathcal{X}} P_d(c_k = x) f_k(c_k = x, \theta_k). \tag{9}$$

We also assume that in the lower part of the FG, describing the phase-noise evolution, a forward–backward node activation schedule is adopted. Therefore, messages $p_f(\theta_k)$ from factor node $p_\Delta(\theta_k - \theta_{k-1})$ to variable node θ_k, and $p_b(\theta_k)$ from factor node $p_\Delta(\theta_{k+1} - \theta_k)$ to variable node θ_k, can be recursively computed as follows:

$$p_f(\theta_k) \propto \int_0^{2\pi} p_d(\theta_{k-1}) p_f(\theta_{k-1}) p_\Delta(\theta_k - \theta_{k-1}) d\theta_{k-1} \tag{10}$$

$$p_b(\theta_k) \propto \int_0^{2\pi} p_d(\theta_{k+1}) p_b(\theta_{k+1}) p_\Delta(\theta_{k+1} - \theta_k) d\theta_{k+1} \tag{11}$$

with uniform pdfs as initial conditions

$$p_f(\theta_0) = p_b(\theta_{K-1}) = \frac{1}{2\pi}. \tag{12}$$

The message $P_u(c_k)$ from f_k to c_k is given by

$$P_u(c_k) \propto \int_0^{2\pi} p_f(\theta_k) p_b(\theta_k) f_k(c_k, \theta_k) d\theta_k. \tag{13}$$

The vector of messages $\{P_u(c_k) : k = 0, \ldots, K-1\}$ represents the observation (in the form of sequence of *a posteriori* pmfs) of the coded symbols "seen" through a virtual memoryless channel, and are processed by the upper part of the graph according to the standard belief propagation algorithm. At each iteration, this produces updated messages $\{P_d(c_k) : k = 0, \ldots, K-1\}$ and updated estimates of the *a posteriori* probabilities $P(b_i|\mathbf{r})$.

Equations (10), (11), and (13), form the main part of the SPA for iterative decoding in the presence of phase noise. It is clear that the implementation complexity of the exact SPA is impractical, since the messages from and to the variable nodes $\{\theta_k\}$ are continuous pdfs. In order to obtain practical algorithms, we follow the canonical distribution approach proposed in [8]. This consists of constraining the messages from/to the continuous variables to take values in a prescribed family of pdfs, that admits a compact parametric representation. Hence, the messages computation reduces to the computation of the pdf parameters. This representation can be exact or, more often, may involve some approximations. In the case of approximations, finding good choices of the pdf parameterization such that the resulting algorithm yields good performance and low computational complexity is generally nontrivial. In the following, we discuss different options for the problem at hand. The proposed algorithms will be compared, in terms of performance and complexity, with the best solutions in the literature that are briefly recalled in Section III.

III. Canonical Distributions Employed in the Literature

A. Discretization of the Channel Parameters: Discretized-Phase BCJR

This case corresponds to letting the canonical distribution be a weighted sum of impulses. This approach has been adopted for Viterbi- and BCJR-like receivers in [24] and [2], [7], [11], and [12], respectively. The channel phase θ_k is assumed to take on the following L values: $\Theta = \{0, 2\pi/L, \ldots, 2\pi(L-1)/L\}$. In [2], the authors found that for M-PSK signals, $L = 8M$ values are sufficient to have no performance loss. This approach is referred to as *discretized-phase BCJR* (dp-BCJR) since after discretization of the phase random-walk the phase trajectories are represented on a trellis diagram with L states, and the SPA message updating rules are identical (not surprisingly) to the BCJR algorithm. Obviously, the dp-BCJR approach becomes "optimal" (in the sense that it approaches the performance of the exact SPA) for a sufficiently large number of discretization levels, at the expenses of an increasing computational complexity.

B. Gaussian Parameterization: Kalman Smoother

Another exemplification of the canonical distribution approach consists of modeling the phasor process $h_k \triangleq e^{j\theta_k}$ as a complex circularly symmetric Gauss–Markov process and treating $\mathbf{h} = (h_0, \ldots, h_{K-1})$ and $\mathbf{r}$ as jointly Gaussian. This assumption yields the forward and backward recursions (10) and (11) in the form of a Kalman smoother [11] (also, see [25, Ch. 5] for a detailed description).

IV. Proposed Algorithms

A. Fourier Parameterization

The function $f_k(c_k, \theta_k)$ defined in (7) is periodic in θ_k. Hence, it can be expanded in Fourier series. We use the well-known identity [26, eq. (9.6.34)]

$$e^{x\cos\theta} = \mathrm{I}_0(x) + 2\sum_{\ell=1}^{\infty} \mathrm{I}_\ell(x)\cos(\ell\theta) \tag{14}$$

where $\mathrm{I}_\ell(x)$ is the modified Bessel function of the first kind of order ℓ. Letting, for a complex number z, $\phi(z) = \arg(z)$, after some straightforward manipulations, we obtain

$$f_k(c_k, \theta_k) \propto e^{-\frac{|c_k|^2}{2\sigma^2}} \sum_{\ell=-\infty}^{\infty} \mathrm{I}_\ell\left(\frac{|r_k||c_k|}{\sigma^2}\right) e^{-j\ell\phi(r_k c_k^*)} e^{j\ell\theta_k}. \tag{15}$$

Substituting (15) into (9), we may express

$$p_d(\theta_k) \propto \sum_{\ell=-\infty}^{\infty} A_k^{(\ell)} e^{j\ell\theta_k} \tag{16}$$

having defined

$$\begin{aligned} A_k^{(\ell)} &\triangleq \sum_{x\in\mathcal{X}} P_d(c_k = x) e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{I}_\ell\left(\frac{|r_k||x|}{\sigma^2}\right) e^{-j\ell\phi(r_k x^*)} \\ &= e^{-j\ell\phi(r_k)} \sum_{x\in\mathcal{X}} P_d(c_k = x) e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{I}_\ell\left(\frac{|r_k||x|}{\sigma^2}\right) \frac{x^\ell}{|x|^\ell}. \end{aligned} \tag{17}$$

Note that for M-PSK signals, the expression of coefficients $A_k^{(\ell)}$, neglecting irrelevant proportionality terms (see footnote 4), simplifies to

$$A_k^{(\ell)} = e^{-j\ell\phi(r_k)} \mathrm{I}_\ell\left(\frac{|r_k|}{\sigma^2}\right) \sum_{x\in\mathcal{X}} P_d(c_k = x)x^\ell. \tag{18}$$

In this case, at the first iteration, when the probabilities of symbols $P_d(c_k)$ are all equal to $1/M$ (except for the pilot symbols), these coefficients are zero for $\ell \neq 0, \pm M, \pm 2M, \pm 3M, \ldots$.

Pdfs $p_f(\theta_k)$ and $p_b(\theta_k)$ take on the same form, i.e., they are periodic as well and can be expanded in Fourier series as

$$p_f(\theta_k) = \sum_{\ell=-\infty}^{\infty} B_{f,k}^{(\ell)} e^{j\ell\theta_k}, \quad p_b(\theta_k) = \sum_{\ell=-\infty}^{\infty} B_{b,k}^{(\ell)} e^{j\ell\theta_k}. \tag{19}$$

Substituting (16) and the above expression for $p_f(\theta_k)$ in (10), we obtain

$$\sum_{\ell=-\infty}^{\infty} B_{f,k}^{(\ell)} e^{j\ell\theta_k} = \sum_{\ell=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} A_{k-1}^{(m)} B_{f,k-1}^{(\ell-m)} \cdot \int_0^{2\pi} e^{j\ell\theta_{k-1}} p_\Delta(\theta_k - \theta_{k-1}) d\theta_{k-1}. \tag{20}$$

We notice that, for practical values of σ_Δ, the pdf $p_\Delta(\varphi)$ is essentially 0 for argument φ outside an interval centered in 0 of size much smaller than 2π. Hence, we can write

$$\begin{aligned}\int_0^{2\pi} & e^{j\ell\theta_{k-1}} p_\Delta(\theta_k - \theta_{k-1}) d\theta_{k-1} \\ &\simeq \int_{-\infty}^{\infty} e^{j\ell\theta_{k-1}} g\left(0, \sigma_\Delta^2, \theta_k - \theta_{k-1}\right) d\theta_{k-1} \\ &= D_\ell(\sigma_\Delta) e^{j\ell\theta_k}\end{aligned} \tag{21}$$

where we define $D_\ell(\sigma_\Delta) \triangleq e^{-(\sigma_\Delta^2 \ell^2/2)}$. By using (21) in (20), we obtain

$$\sum_{\ell=-\infty}^{\infty} B_{f,k}^{(\ell)} e^{j\ell\theta_k} = \sum_{\ell=-\infty}^{\infty} \left[D_\ell(\sigma_\Delta) \sum_{m=-\infty}^{\infty} A_{k-1}^{(m)} B_{f,k-1}^{(\ell-m)}\right] e^{j\ell\theta_k} \tag{22}$$

yielding the forward recursion for the Fourier coefficients $B_{f,k}^{(\ell)}$

$$B_{f,k}^{(\ell)} = D_\ell(\sigma_\Delta) \sum_{m=-\infty}^{\infty} A_{k-1}^{(m)} B_{f,k-1}^{(\ell-m)} = D_\ell(\sigma_\Delta)\left[A_{k-1}^{(\ell)} \otimes B_{f,k-1}^{(\ell)}\right] \tag{23}$$

where $\otimes$ denotes convolution of sequences. From condition (12), we derive the initial condition $B_{f,0}^{(\ell)} = \delta(\ell)$, where $\delta(\ell)$ denotes the Kronecker delta. Similarly, the backward recursion to compute the coefficients $\{B_{b,k}^{(\ell)}\}$ is given by

$$B_{b,k}^{(\ell)} = D_\ell(\sigma_\Delta)\left[A_{k+1}^{(\ell)} \otimes B_{b,k+1}^{(\ell)}\right] \tag{24}$$

with initial condition $B_{b,K-1}^{(\ell)} = \delta(\ell)$.

The computation of these coefficients can be simplified taking into account the symmetries $A_k^{(-\ell)} = A_k^{(\ell)*}$, $B_{f,k}^{(-\ell)} = B_{f,k}^{(\ell)*}$, and $B_{b,k}^{(-\ell)} = B_{b,k}^{(\ell)*}$. Finally, substituting (15) and (19) into (13) and defining

$$E_k^{(\ell)} \triangleq e^{-\frac{|c_k|^2}{2\sigma^2}} \left\{B_{f,k}^{(\ell)} \otimes B_{b,k}^{(\ell)} \otimes \left[\mathrm{I}_\ell\left(\frac{|r_k||c_k|}{\sigma^2}\right) e^{-j\ell\phi(r_k c_k^*)}\right]\right\} \tag{25}$$

we have

$$P_u(c_k) \propto \sum_{\ell=-\infty}^{\infty} E_k^{(\ell)} \int_0^{2\pi} e^{j\ell\theta_k} d\theta_k = E_k^{(0)}. \tag{26}$$

Remark: Truncation of the Fourier Coefficients: The convolution of the infinite-length sequence of Fourier coefficients can be effectively implemented by truncation. Hence, only a reduced number N of coefficients can be taken into account due to the fact that, for a given x, functions $\mathrm{I}_\ell(x)$ are monotonically decreasing for increasing values of ℓ. Standard smoothed truncation methods (windowing) can be applied [27]. In particular, we found experimentally that a Kaiser window with optimized parameter β [27] yields satisfactory results, as it will be demonstrated in Section VI.

B. Tikhonov Parameterization

Let us consider (9). If the messages $P_d(c_k)$ were the exact probabilities of the code symbols, it would be

$$p_d(\theta_k) \propto \sum_{x\in\mathcal{X}} P_d(c_k = x) f_k(c_k = x, \theta_k) \propto p(r_k|\theta_k). \tag{27}$$

We approximate $p(r_k|\theta_k)$ by the nearest Gaussian pdf in the sense of divergence (Kullbach–Leibler distance) [28]. This yields the Gaussian pdf with mean $\mathbb{E}[r_k|\theta_k]$ and variance $\mathrm{var}(r_k|\theta_k)$. Letting α_k and β_k be the first- and second-order moments of $c_k \sim P_d(c_k)$, given by

$$\alpha_k \triangleq \sum_{x\in\mathcal{X}} x P_d(c_k = x), \quad \beta_k \triangleq \sum_{x\in\mathcal{X}} |x|^2 P_d(c_k = x) \tag{28}$$

we obtain

$$\begin{aligned} p(r_k|\theta_k) &\simeq g_{\mathbb{C}}\left(\alpha_k e^{j\theta_k}, 2\sigma^2 + \beta_k - |\alpha_k|^2, r_k\right) \\ &\propto \exp\left\{2\frac{\mathrm{Re}\left[r_k \alpha_k^* e^{-j\theta_k}\right]}{2\sigma^2 + \beta_k - |\alpha_k|^2}\right\}. \end{aligned} \tag{29}$$

Substituting (29) in the forward recursion (10), we obtain

$$p_f(\theta_k) \simeq \int_0^{2\pi} \exp\left\{2\frac{\mathrm{Re}\left[r_{k-1}\alpha_{k-1}^* e^{-j\theta_{k-1}}\right]}{2\sigma^2 + \beta_{k-1} - |\alpha_{k-1}|^2}\right\} \cdot p_f(\theta_{k-1}) p_\Delta(\theta_k - \theta_{k-1}) d\theta_{k-1}. \tag{30}$$

When the channel phase is slowly varying, i.e., for $\sigma_\Delta \to 0$, we have $p_\Delta(\theta_k - \theta_{k-1}) = \delta(\theta_k - \theta_{k-1})$. In this case, the solution of the recursion given by (30) with initial condition (12) is a sequence of Tikhonov pdfs, given by

$$p_f(\theta_k) \propto \exp\left\{\mathrm{Re}[a_{f,k} e^{-j\theta_k}]\right\} \tag{31}$$

TABLE I
COMPUTATIONAL LOAD PER CODE SYMBOL PER ITERATION FOR M-PSK MODULATIONS

	Tikhonov	Kalman	Fourier	dp-BCJR
Operations	$17M+11$	$67M+53$	$(\frac{M}{2}+4)N^2+(2M-9)N-2M+8$	$13ML+10QL-9L-3M$
LUT accesses	$3M+3$	$17M+4$	$2M+2(N+1)$	$3ML+2QL-3L-M$

where $a_{f,k}$ can be recursively computed as

$$a_{f,k}=a_{f,k-1}+2\frac{r_{k-1}\alpha_{k-1}^*}{2\sigma^2+\beta_{k-1}-|\alpha_{k-1}|^2} \quad (32)$$

with the initial condition $a_{f,0}=0$. Similarly, the solution of the backward recursion (11) under the above approximations is the sequence of Tikhonov pdfs

$$p_b(\theta_k)\propto \exp\left\{\mathrm{Re}[a_{b,k}e^{-j\theta_k}]\right\} \quad (33)$$

where $a_{b,k}$ can be recursively computed as

$$a_{b,k}=a_{b,k+1}+2\frac{r_{k+1}\alpha_{k+1}^*}{2\sigma^2+\beta_{k+1}-|\alpha_{k+1}|^2} \quad (34)$$

with the initial condition $a_{b,K-1}=0$. From (13), (31), and (33), we obtain

$$P_u(c_k)\propto \exp\left\{-\frac{|c_k|^2}{2\sigma^2}\right\}\mathrm{I}_0\left(\left|a_{f,k}+a_{b,k}+\frac{r_kc_k^*}{\sigma^2}\right|\right). \quad (35)$$

When the phase varies more rapidly, such that the approximation $p_\Delta(\theta_k-\theta_{k-1})\simeq\delta(\theta_k-\theta_{k-1})$ is no longer valid, we show in Appendix A that the distributions $p_f(\theta_k)$ and $p_b(\theta_k)$ are still approximately given in the form (31) and (33), where now the coefficients $a_{f,k}$ and $a_{b,k}$ are updated by the modified forward and backward recursions

$$a_{f,k}=\left[a_{f,k-1}+2\frac{r_{k-1}\alpha_{k-1}^*}{2\sigma^2+\beta_{k-1}-|\alpha_{k-1}|^2}\right]\cdot\gamma\left(\sigma_\Delta^2,\left|a_{f,k-1}+2\frac{r_{k-1}\alpha_{k-1}^*}{2\sigma^2+\beta_{k-1}-|\alpha_{k-1}|^2}\right|\right) \quad (36)$$

$$a_{b,k}=\left[a_{b,k+1}+2\frac{r_{k+1}\alpha_{k+1}^*}{2\sigma^2+\beta_{k+1}-|\alpha_{k+1}|^2}\right]\cdot\gamma\left(\sigma_\Delta^2,\left|a_{f,k+1}+2\frac{r_{k+1}\alpha_{k+1}^*}{2\sigma^2+\beta_{k+1}-|\alpha_{k+1}|^2}\right|\right) \quad (37)$$

where we define the function

$$\gamma(x_1,x_2)=\frac{1}{1+x_1x_2}. \quad (38)$$

In a practical implementation, the function $\gamma(x_1,x_2)$ can be computed via a lookup table (LUT).

Remark: Modification in the Case of Long Pilot Fields: When the pilot symbols are arranged in bursts (training sequences) separated by long blocks of code symbols, as in the case of the DVB-S2 system [29], it is necessary to slightly modify the algorithm in order to speedup the convergence process and to avoid the risk of a phase ambiguity. In fact, consider the recursive integral equation (10) from the second iteration on. If the product

$$p_d(\theta_k)p_f(\theta_k)=\left(\sum_{x\in\mathcal{X}}P_d(c_k=x)f_k(c_k=x,\theta_k)\right)p_f(\theta_k)$$

TABLE II
COMPUTATIONAL LOAD PER CODE SYMBOL PER ITERATION FOR $M=4$, $L=8$, $M=32$, $N=17$, AND $Q=3$

	Tikhonov	Kalman	Fourier	dp-BCJR
Operations	79	321	1717	2324
LUT accesses	15	72	44	476

contains a dominant exponential term, i.e., if there exists a modulation symbol $\overline{x}\in\chi$ such that

$$\ln P_d(c_k=\overline{x})+\left|a_{f,k}+\frac{r_k\overline{x}^*}{\sigma^2}\right|>\delta+\ln P_d(c_k=x)+\left|a_{f,k}+\frac{r_kx^*}{\sigma^2}\right|,\quad \forall x\in\mathcal{X},x\neq\overline{x} \quad (39)$$

where $\delta>0$ is a real parameter to be optimized by computer simulation, it is preferable to let $\alpha_k=\overline{x}$ and $\beta_k=|\overline{x}|^2$. Otherwise, we choose α_k and β_k, as in (28). This corresponds to using a decision-aided scheme based on hard decisions for some symbols c_k. Similar considerations also hold for the recursive integral equation (11). In the numerical results related to the DVB-S2 system, we found that $\delta=1.5$ yields satisfactory results.

V. COMPLEXITY CONSIDERATIONS

We address the computational complexity of the proposed algorithms and compare it with the complexity of the dp-BCJR and the Kalman smoother. We assume that the computation of non linear functions is performed by using a lookup table and restrict our evaluation to the case of M-PSK. Table I presents computational complexity in terms of number of operations (additions and multiplications) between two real arguments and accesses to LUT, per coded symbols per decoder iteration. For the dp-BCJR algorithm, the integer parameter Q has been further introduced. In fact, being Δ_k a Gaussian random variable with small variance, only the transitions from any phase state to the $Q<L$ adjacent phase states can be taken into account in the resulting trellis. For example, in [2], $Q=3$ is used, meaning that the phase can either remain constant or change by $\pm2\pi/L$ at every kth trellis step. It should be noticed that the value of Q depends on the number of discretization levels L, i.e., ultimately on the modulation constellation size M (see [2] and Section III-A).

As an example, the number of operations per code symbol per iterations for quadrature phase-shift keying (QPSK) ($M=4$), $L=8M=32$, $N=17$, and $Q=3$ is reported in Table II. The complexity advantage of the algorithm based on the Tikhonov parameterization is clear. It should also be noticed that, depending on the implementation, complexity should be measured in different ways. For example, for a DSP implementation additions and multiplications have the same cost, while for a very large scale integration (VLSI) implementation the chip area and the suitability to parallel computation should also be taken into account. Such detailed evaluation, specific to a given implementation, is beyond the scope of this paper.

VI. Numerical Results

In this section, the performance of the proposed schemes 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 $N_0 = \sigma^2$ the one-sided noise power spectral density. Unless otherwise stated, the considered code is a (3,6)-regular LDPC code of length 4000 [30] with binary phase-shift keying (BPSK) modulation and a maximum of 200 iterations of the SPA on the overall graph is allowed. For each simulated point, a minimum of 100 frame errors were counted.

In all simulated cases, pilot symbols are inserted in the transmitted codeword in order to make the iterative decoding algorithms bootstrap. In fact, by studying the extrinsic information transfer (EXIT) charts [31] of the overall detector/decoder, not shown here for the sake of space limitations,[5] it can be observed that the iterative decoding system has a fixed point at zero extrinsic information, irrespectively of the SPA approximation adopted. This means that in the absence of pilot symbols the iterative decoder will be stuck at zero symbol reliability forever. This observation clearly identifies the role of pilot symbols in iterative joint decoding and phase estimation: they are analogous of "doping" symbols currently used in turbo-coding design, in order to remove the zero fixed point and make the iterative decoder bootstrap (see, for example, [31] and [33]).

Pilot symbols involve a slight decrease of the effective information rate, resulting in an increase in the required signal-to-noise ratio. This increase has been introduced artificially in the curve labeled "known phase" for the sake of comparison. Hence, the gap between the "known phase" curve and the others is uniquely due to the need for phase estimation/compensation, and not to the rate decrease due to pilot symbols.

Beyond the Wiener phase noise model described in Section II, we considered the DVB-S2 compliant ESA phase noise model given as follows: $\{\theta_k\}$ is the sum of the outputs of two infinite impulse response filters driven by the same white Gaussian noise process with unit variance, where the filters are chosen to fit an experimental phase noise mask. The filter transfer functions are given by (see [22] and [23] for details)

$$H_1(z)=\frac{1}{\sqrt{2T}}\frac{-4.7\cdot 10^{-11}}{(z-0.999\,975)^2}$$
$$H_2(z)=\frac{1}{\sqrt{2T}}\frac{2.8\cdot 10^{-6}(z-0.992\,015)(z-1.103\,181)}{(z-0.991\,725)(z-0.999\,998\,5)(z-0.563\,507)}$$

where T is the symbol interval.

In Fig. 3, the newly proposed algorithm based on Fourier parameterization is compared with the dp-BCJR (our benchmark algorithm). One pilot symbol in every block of 20 transmitted symbols was used. The ESA phase noise model and a more severe Wiener model (2) with $\sigma_\Delta = 6°$, have been considered. In the case of the ESA model, all the receivers were designed by assuming a Wiener phase noise model with $\sigma_\Delta = 0.3°$, optimized via simulation.

In the case of the Wiener model, different values L of discretization levels and different values N of the Fourier coefficients have been considered. No improvement has been observed for values of $L > 16$ and this is in agreement with a result in [2]. Similarly, values of $N > 17$ were not considered since they do not produce any performance improvement. Therefore, the value of $N = 17$ (i.e., $-8 \le \ell \le 8$ in all the equations of Section IV-A) can be considered as nearly-optimal for $\sigma_\Delta = 6°$.

The advantage of the Fourier algorithm over the dp-BCJR appears for high-order modulations. In fact, while the dp-BCJR requires a number of states that increases linearly with the modulation size, the new algorithm needs a nearly constant number of Fourier coefficients. This fact is highlighted in Fig. 4, where the QPSK modulation is considered. Again, a Wiener phase noise with $\sigma_\Delta = 6°$ and one pilot symbol every 20 transmitted symbols is used. The dp-BCJR algorithm is applied with $L = 16$ and with $L = 8M = 32$ quantization levels (phase states), whereas

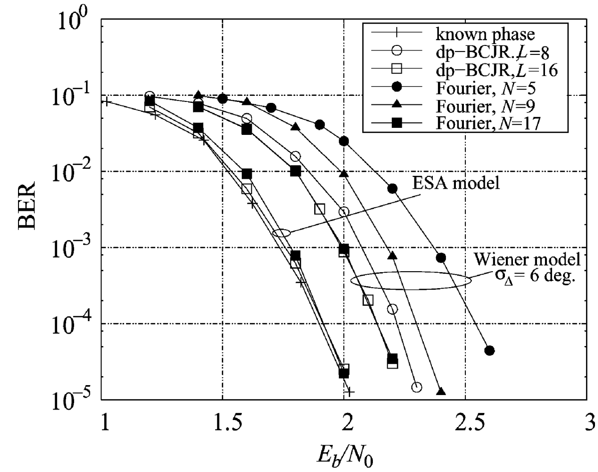

Fig. 3. Performance of the algorithms based on discretization of channel parameters (dp-BCJR) and Fourier parameterization. BPSK and two different phase models are considered.

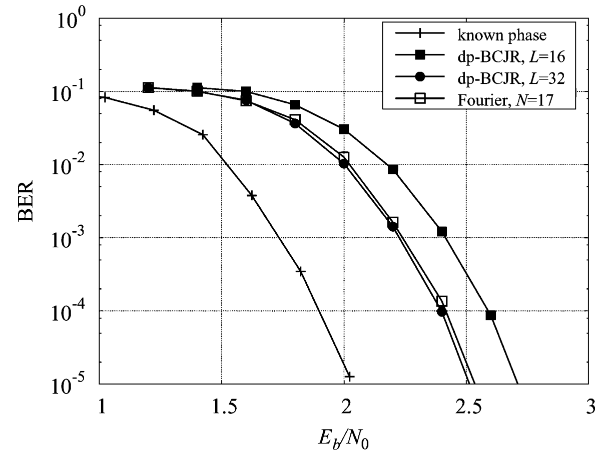

Fig. 4. Performance of the algorithms based on discretization of channel parameters (dp-BCJR) and Fourier parameterization. QPSK and the Wiener model with $\sigma_\Delta = 6°$ are considered.

[5]The interested reader may refer to [32] for the relevant curves.

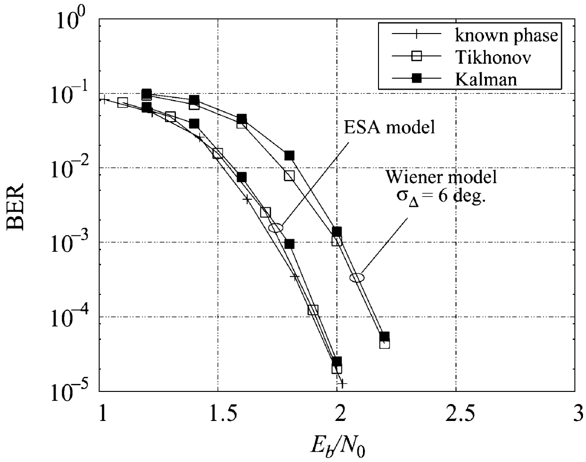

Fig. 5. Performance of the algorithm based on Tikhonov parameterization and the Kalman smoother. BPSK and two different phase models are considered.

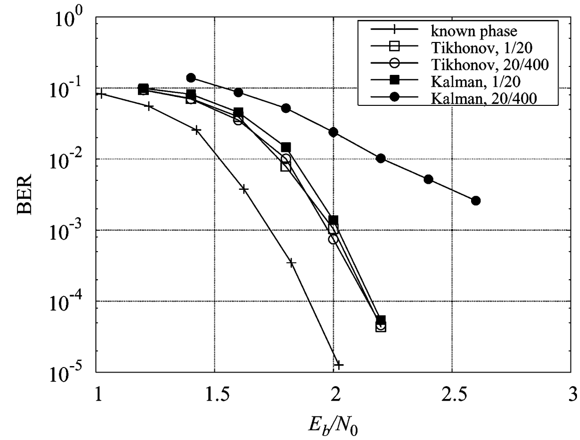

Fig. 7. Performance of the algorithms based on Tikhonov parameterization and the Kalman smoother. BPSK and two different pilot distributions are considered.

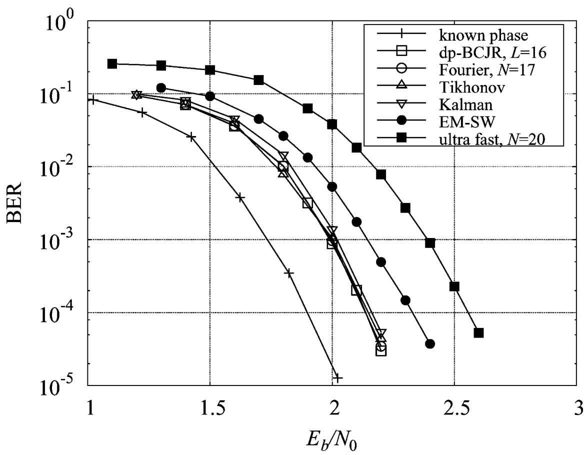

Fig. 6. Performance of all the proposed algorithms and comparison with other algorithms proposed in the literature. BPSK and the Wiener phase model with $\sigma_\Delta = 6^\circ$ are considered.

the Fourier algorithm makes use of $N = 17$ coefficients. We notice that the dp-BCJR needs $L = 32$ states, in agreement with the findings of [2], yielding practically the same performance of the Fourier algorithm with only $N = 17$, with evident complexity savings (see Table II).

In Fig. 5, the newly proposed algorithm based on Tikhonov parameterization is compared with the Kalman smoother in the same conditions of Fig. 3. We observe that, despite the very low complexity, both algorithms have practically the same performance of the much more computationally demanding dp-BCJR and Fourier algorithms. This fact can be also observed from Fig. 6, where all the considered algorithms are compared for the Wiener phase noise with $\sigma_\Delta = 6^\circ$. In this figure, the performance of two other algorithms described in the literature is also shown for the sake of comparison. The first one is the "ultrafast" algorithm with overlapped windows described in [21], with the value of N optimized by computer simulation. The second one is based on the EM algorithm [15]–[19]. In order to adapt the algorithm to a time-varying channel phase, a sliding-window version of the EM algorithm is used, where the window size was optimized by computer simulation. The resulting algorithm is denoted by sliding-window EM (EM-SW). We found that the optimal window has width of 60 symbols for the considered phase noise. In both cases, the performance loss is due to the fact that these two algorithms are designed for a different phase model, i.e., a block-constant phase. Based on the above experiments and on extensive numerical evidence (not shown for the sake of space limitation), we conclude that all the proposed algorithms exhibit a practically optimal performance (i.e., they perform as well as the discretization approach). Among them, the algorithm based on the Tikhonov parameterization is particularly attractive because of its low complexity.

It remains to assess the advantage of this latter algorithm with respect to the Kalman smoother, of roughly the same complexity (see Table I). This is evidenced in terms of the sensitivity to the placement of pilot symbols. Fig. 7 shows the performance for the Wiener model with $\sigma_\Delta = 6^\circ$ and two different distributions, namely, 1 pilot symbol in each block of 20 consecutive transmitted symbols and 20 pilots in each block of 400 consecutive transmitted symbols, such that the effective information rate is the same. We may observe that the algorithm based on Tikhonov parameterization is almost insensitive thanks to the algorithm modification described in Section IV-B. A similar modification is not possible in the case of the Kalman smoother, since it can be shown that the choice of a dominant term corresponds to a hard-decision based uniquely on the decoder outcome $P_d(c_k)$. We verified that this modification produces no performance improvement for the Kalman smoother. We interpret this fact by noticing that the Tikhonov parameterization makes explicit use of the fact that the phasor $e^{j\theta_k}$ must lie on the unit circle, while the Kalman smoother, assuming that the phasor is a Gauss–Markov process, is somehow more mismatched with respect to the true statistics of the channel.

Note that, in general, the distribution of the pilots has to be optimized for the specific detection algorithm employed. Nevertheless, in many communications standards the placement of pilot symbols is determined *a priori*, without any specific detection algorithm in mind. Therefore, an algorithm which is almost insensitive to the placement of pilot symbols is very useful in practice. For example, this is the case of the DVB-S2 system, where pilot symbols are organized into bursts of 36 symbols every 1476 transmitted symbols [29]. We consider two standard-

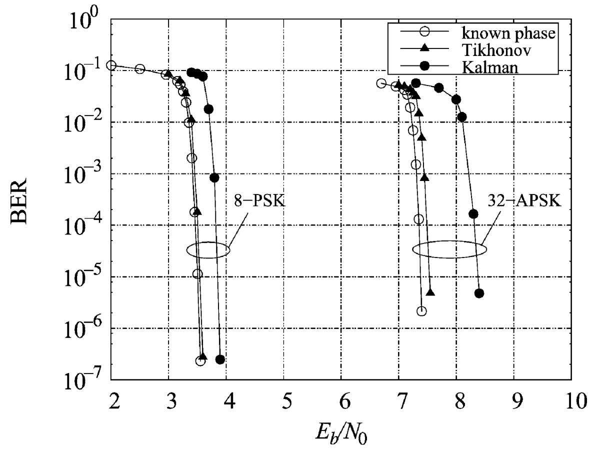

Fig. 8. Performance of the algorithms based on Tikhonov parameterization and the Kalman smoother. The ESA phase model is considered along with 8-PSK and 32-APSK modulations.

ized LDPC codes with codewords of length 64 800 [29]. The first one has rate 2/3 and is mapped onto an 8-PSK modulation. The second one has rate 4/5 and is mapped onto a 32-APSK modulation. The above mentioned phase noise ESA model is considered. The performance is shown in Fig. 8. For the algorithm based on Tikhonov parameterization, the loss due to phase noise is less than 0.1 dB in both cases. Notice that a further improvement in performance may be obtained if the maximum number of iterations is not limited to 50. The Kalman smoother does not perform as well mainly because of the bursty allocation of pilot symbols.

VII. Conclusion

In this paper, the problem of iterative decoding in AWGN channels affected by phase noise has been considered. We proposed two new algorithms based on suitable approximations of the SPA applied to the FG representing the posterior joint pmf of the information bits and the random channel phase process given the received signal. Our approximations are based on the concept of *canonical distributions*, i.e., we impose that the messages propagated by the SPA take on values in a certain parametric family of distribution functions, such that the approximated SPA reduces to updating and propagating the function parameters. We proposed two new parameterizations: one based on Fourier series and the other based on a minimum divergence Gaussian approximation and on the Tikhonov distribution.

Among the considered schemes, the novel algorithm based on Tikhonov parameterization exhibits practically optimal performance and very low complexity, and represents an attractive solution for systems, where powerful LDPC-coded modulations are transmitted in the presence of phase noise with bursty pilot symbols, such as in next-generation satellite DVB.

While in this paper we have considered the case of perfect frequency synchronization, an important extension of this work should consider also frequency errors. After the first submission of this paper, we extended the Tikhonov parameterization algorithm to the case of channels affected by a random but constant frequency offset uniformly distributed in an interval centered around the nominal carrier frequency [34].

Appendix
Modified Tikhonov Parameterization

The integral (30), whose domain can be any interval of length 2π, when $p_f(\theta_{k-1})$ is in the form (31) and $p_\Delta(\theta_k - \theta_{k-1}) \simeq g(0, \sigma_\Delta^2, \theta_k - \theta_{k-1})$ can be put in the form

$$f(y,z) = \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \int_{y-\pi}^{y+\pi} e^{\mathrm{Re}[ze^{-jx}]} e^{-\frac{(x-y)^2}{2\sigma_\Delta^2}}\, dx \tag{40}$$

for a suitable choice of the complex parameter z and real variables x and y. By discarding irrelevant multiplicative factors, we shall show that $f(y,z) \simeq e^{\gamma(\sigma_\Delta^2, |z|)\mathrm{Re}[ze^{-jy}]}$, where $\gamma(x_1, x_2)$ is given in (38). To this purpose, we use the following approximation which holds for large values of $a \in \mathbb{R}^+$ (in practice, $a > 5$)

$$\begin{aligned}\frac{e^{a\cos(x-y)}}{2\pi \mathrm{I}_0(a)} &\simeq \frac{1}{\sqrt{\frac{2\pi}{a}}} e^{-\frac{a}{2}(x-y)^2}\\ &= g\left(y, \frac{1}{a}, x\right), \quad \text{for } y-\pi \le x \le y+\pi. \end{aligned}\tag{41}$$

In fact, for sufficiently large values of a, the Tikhonov pdf $e^{a\cos(x-y)}/2\pi \mathrm{I}_0(a)$ has its support in a small interval around y. Hence, by using a second-order Taylor expansion, we have $\cos(x-y) \simeq 1 - ((x-y)^2)/2$. A normalization constant has been further added to obtain a pdf. Then, using (41) in (40), we obtain

$$\begin{aligned} f(y,z) &\overset{(a)}{\simeq} \int_{-\infty}^{\infty} e^{\mathrm{Re}[ze^{-jx}]} g\left(x, \sigma_\Delta^2, y\right) dx\\ &\overset{(b)}{\simeq} 2\pi \mathrm{I}_0\left(|z|\right) \int_{-\infty}^{\infty} g\left(\phi(z), \frac{1}{|z|}, x\right) g\left(x, \sigma_\Delta^2, y\right) dx\\ &\overset{(c)}{\propto} 2\pi \mathrm{I}_0\left(|z|\right) g\left(\phi(z), \frac{1}{|z|} + \sigma_\Delta^2, y\right)\\ &\overset{(d)}{\simeq} \frac{\mathrm{I}_0\left(|z|\right)}{\mathrm{I}_0\left(\frac{|z|}{1+\sigma_\Delta^2|z|}\right)} \exp\left\{\frac{1}{1+\sigma_\Delta^2|z|}\mathrm{Re}[ze^{-jy}]\right\}\\ &\propto \exp\left\{\frac{1}{1+\sigma_\Delta^2|z|}\mathrm{Re}[ze^{-jy}]\right\} \end{aligned}\tag{42}$$

where (a) follows from the observation that, for $\sigma_\Delta \ll 2\pi$, the function $e^{-((x-y)^2)/2\sigma_\Delta^2}$ has its support in a small interval around y, (b) and (d) follow from (41) and (c) can be easily proved by direct calculation. Eventually, we obtain

$$\gamma\left(\sigma_\Delta^2, |z|\right) = \frac{1}{1+\sigma_\Delta^2|z|}.$$

References

[1] 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.

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

[3] I. D. Marsland and P. T. Mathiopoulos, "On the performance of iterative noncoherent detection of coded M-PSK signals," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 588–596, Apr. 2000.

[4] G. Colavolpe, G. Ferrari, and R. Raheli, "Noncoherent iterative (turbo) detection," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1488–1498, Sep. 2000.
[5] G. Ferrari, G. Colavolpe, and R. Raheli, "Noncoherent iterative decoding of spectrally efficient coded modulations," *Ann. Telecommun.*, vol. 56, pp. 409–421, Jul./Aug. 2001.
[6] A. Anastasopoulos and K. M. Chugg, "Adaptive iterative detection for phase tracking in turbo coded systems," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 2135–2144, Dec. 2001.
[7] 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.
[8] 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.
[9] H. Jin and T. Richardson, "Design of low-density parity-check codes for noncoherent MPSK communication," in *Proc. IEEE Symp. Inf. Theory*, Lausanne, Switzerland, Jun.–Jul. 2002, p. 169.
[10] ——, "On iterative joint decoding and demodulation," in *Proc. 41st Allerton Conf.*, Monticello, IL, Oct. 2003.
[11] J. Dauwels and H.-A. Loeliger, "Joint decoding and phase estimation: An exercise in factor graphs," in *Proc. IEEE Symp. Inf. Theory*, Yokohama, Japan, Jul. 2003, p. 231.
[12] ——, "Phase estimation by message passing," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, Jun. 2004, pp. 523–527.
[13] 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.
[14] G. Colavolpe, "LDPC codes over channels with memory," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, Jun. 2004, pp. 757–761.
[15] V. Lottici and M. Luise, "Carrier phase recovery for turbo-coded linear modulations," in *Proc. IEEE Int. Conf. Commun.*, New York, Apr. 2002, pp. 1541–1545.
[16] ——, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, Apr. 2004.
[17] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo synchronization: An EM algorithm interpretation," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, Jun. 2003, pp. 2933–2937.
[18] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft information based synchronization in iterative (turbo) receivers," IEEE Trans. Commun., 2003, submitted for publication.
[19] H. Steendam, N. Noels, and M. Moeneclaey, "Iterative carrier phase synchronization for low-density parity-check coded systems," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, Jun. 2003, pp. 3120–3124.
[20] L. Zhang and A. Burr, "Application of turbo principle to carrier phase recovery in turbo encoded bit-interleaved coded modulation system," in *Proc. Int. Symp. Turbo Codes, Relat. Topics*, Brest, France, Sep. 2003, pp. 87–90.
[21] 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.
[22] A. Ginesi, D. Fittipaldi, A. Bigi, and R. De Gaudenzi, "Pilot-aided carrier synchronization techniques for broadband satellite transmissions," ESA-ESTEC, Tech. Rep., Sep. 2003.
[23] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," in *Proc. 8th Int. Workshop Signal Process. Space Commun.*, Catania, Italy, Sep. 2003, pp. 177–184.
[24] O. Macchi and L. L. Scharf, "A dynamic programming algorithm for phase estimation and data decoding on random phase channels," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 581–595, Sep. 1981.
[25] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection Algorithms for Wireless Communications*. New York: Wiley, 2004.
[26] *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, Eds., Dover, New York, 1972.
[27] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[29] ETSI, "Digital video broadcasting (DVB): Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications," ETSI, DVBS2 74r13, 2003.
[30] D. J. C. MacKay. Regular LDPC online database. [Online]. Available: http://www.inference.phy.cam.ac.uk/mackay/
[31] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEU Int. J. Electronic. Commun.*, vol. 54, no. 6, pp. 389–398, Dec. 2000.
[32] G. Colavolpe and G. Caire, "Iterative joint detection and decoding for communications under random time-varying carrier phase," Tech. Rep., ESA Contract 17337/03/NL/LvH, Mar. 2004.
[33] S. ten Brink, "Code doping for triggering iterative decoding convergence," in *Proc. IEEE Symp. Inf. Theory*, Jun. 2001, p. 235.
[34] A. Barbieri, G. Colavolpe, and G. Caire, "Joint iterative detection and decoding in the presence of phase noise and frequency offset," in *Proc. IEEE Int. Conf. Commun.*, May 2005.

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

Since 1997, he has been at the University of Parma, where he is now an Associate Professor of Telecommunications. In 2000, he was a Visiting Scientist at the Institut Eurécom, Valbonne, France. He is a coauthor of *Detection Algorithms for Wireless Communications* (New York: Wiley, 2004). His research activity has led to numerous scientific publications in leading international journals and conference proceedings and a few industrial patents. His main research interests include digital transmission theory, channel coding, and signal processing.

Alan Barbieri (S'05) was born in Parma, Italy, in 1979. He received the Dr. Ing. degree in telecommunications engineering (*cum laude*) from the University of Parma, Parma, Italy, in 2003. Currently, he is working towards the Ph.D. degree at the Dipartimento di Ingegneria dell'Informazione (DII), University of Parma.

His main research interests include channel coding, iterative joint detection and decoding algorithms, estimation of unknown parameters, and algorithms for synchronization.

Giuseppe Caire (S'92–M'94–SM'03–F'05) was born in Torino, Italy, in 1965. He received the B.Sc. degree in electrical engineering from Politecnico di Torino, Torino, Italy, in 1990, the M.Sc. degree in electrical engineering from Princeton University, Princeton, NJ, in 1992, and the Ph.D. degree from Politecnico di Torino in 1994.

He was a recipient of the AEI G. Someda Scholarship in 1991, the COTRAO Scholarship in 1996, and of a CNR Scholarship in 1997. He was with the European Space Agency (ESTEC), Noordwijk, The Netherlands, from May 1994 to February 1995. He visited Princeton University in Summer 1997 and Sydney University in Summer 2000. He has been an Assistant Professor of Telecommunications at the Politecnico di Torino and presently is a Professor with the Department of Mobile Communications, Institute Eurécom. His current interests are in the field of communications theory, information theory, and coding theory with particular focus on wireless applications.

Dr. Caire received the Jack Neubauer Best System Paper Award from the IEEE Vehicular Technology Society in 2003, and the Joint IT/Comsoc Best Paper Award in 2004. He served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1998 to 2001, and an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2001 to 2003.