Code-aided Maximum-likelihood Ambiguity Resolution Through Free-energy Minimization

Cedric Herzet, Woradit Kampol, Henk Wymeersch, Luc Vandendorpe

Code-aided Maximum-likelihood Ambiguity Resolution Through Free-energy Minimization

Cédric Herzet, *Member, IEEE*, Kampol Woradit, *Student Member, IEEE*,
Henk Wymeersch, *Member, IEEE*, and Luc Vandendorpe, *Fellow, IEEE*

Abstract

In digital communication receivers, ambiguities in terms of timing and phase need to be resolved prior to data detection. In the presence of powerful error-correcting codes, which operate in low signal to noise ratios (SNR), long training sequences are needed to achieve good performance. In this contribution, we develop a new class of code-aided ambiguity resolution algorithms, which require no training sequence and achieve good performance with reasonable complexity. In particular, we focus on algorithms that compute the maximum-likelihood (ML) solution (exactly or in good approximation) with a tractable complexity, using a factor-graph representation. The complexity of the proposed algorithm is discussed, and reduced complexity variations, including stopping criteria and sequential implementation, are developed.

I. Introduction

The problem of timing and phase ambiguity resolution is encountered in most digital communication receivers. Resolving such ambiguities is of major importance, since incorrect resolution usually leads to the loss of the entire data packet. Conventional methods dealing with ambiguity resolution operate in Data-aided (DA) mode, i.e., making use of pilot symbols to make a decision. We mention the contributions

This research was partly funded by the network of excellence Newcom++. C. Herzet is with INRIA Centre Rennes - Bretagne Atlantique, Campus universitaire de Beaulieu, Rennes, France (e-mail: cedric.herzet@irisa.fr). K. Woradit is with the Electrical Engineering Department, Faculty of Engineering, Srinakharinwirot University, Nakonnayok, Thailand (e-mail: kampol@swu.ac.th). H. Wymeersch is with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden (e-mail: henk.wymeersch@ieee.org). L. Vandendorpe is with the Communications Laboratory, Université catholique de Louvain, Pl. du Levant 2, B1348 Louvain-la-Neuve, Belgium (e-mail: Luc.Vandendorpe@uclouvain.be).

of Massey [1] and Lui *et al.* [2] on timing ambiguity resolution,[1] and Cacciamani *et al.* [3] on phase ambiguity resolution, in which maximum-likelihood (ML) DA algorithms are derived.

Since the advent of turbo codes [4], [5] in the 90's, ambiguity resolution, and synchronization in general, have become a quite challenging task: on the one hand, synchronizers have to face the extremely low signal-to-noise ratio (SNR) at which these powerful codes operate; on the other hand, the very low bit error rates achieved by such codes imply a high sensitivity to knowledge of the synchronization parameters. In this context, conventional DA ambiguity-resolution methods may require a huge number of pilots to properly synchronize the system, thus leading to an important waste in terms of spectral and power efficiency.

In order to deal with this problem, so-called code-aided (CA) synchronization methods have been proposed in the technical literature, see *e.g.*, [6]–[22]. The idea of CA synchronization is to take benefit from the knowledge of the code structure to improve the estimation quality. The algorithms dealing with the estimation of the fractional part of the synchronization parameters are often referred to as *turbo synchronizers*. In agreement with the turbo principle, these algorithms are based on the exchange of some "soft" information between a synchronization and a detection device. We refer the reader to the following contributions dealing with this problem [6]–[14].

Among CA ambiguity-resolution methods, one can distinguish between two main approaches: the authors either propose ad-hoc algorithms based on the fact that some decoder metrics vary as a function of the considered hypotheses [15]–[19], or place the CA ambiguity-resolution problem in the context of maximum a posteriori (MAP) or ML estimation, [20]–[22]. In the latter class of algorithms, [20] modifies the likelihood function by only keeping its largest term and decides whether the current estimate is true or not by a threshold decision on this modified likelihood function. Another approach is followed in [21]: the authors place the MAP estimation problem into the framework of the expectation-maximization (EM) algorithm [23]. Since the EM algorithm is not suited to discrete estimation problems, the authors propose some judicious ad-hoc modifications which are shown to give very good results in practice. This approach has been given a more rigorous justification in [24]. Finally, in [22] the authors make their decision by maximizing a modified likelihood function, built by using the extrinsic probabilities delivered by the decoder as symbol a priori probabilities.

In this paper, we place the ambiguity-resolution problem within the general framework of free-energy minimization [25]. This approach allows for the tractable approximation of a probability by solving an

[1]Timing ambiguity resolution is often referred to as frame synchronization.

optimization problem involving a (so-called) "free energy" cost function. Inference based on free-energy minimization has already been considered in a number of contributions. In [26] and [27], the authors considered the mean-field approximation of the Gibbs-Helmoltz free energy to estimate respectively the carrier-phase offset and the channel impulse response. A similar approach was taken in [28], [29] for devising OFDM and multi-user receivers with additional constraints (*e.g.*, Gaussianity) on the sought probabilities. Algorithms based on the minimization of the constrained Bethe free energy have also been proposed in the literature. We refer the reader to the following contributions [30]–[32].

In this paper, we propose a new ambiguity resolution method based on the minimization of the constrained Bethe free energy of the system. When the factor graph is cycle-free, the proposed method computes the *exact* MAP (or ML) solution. Moreover, in such a case, the complexity is roughly *half* the complexity of CA methods previously proposed in the literature. In order to further decrease the complexity of the proposed CA ambiguity resolution method, we develop an early stopping rule, as well as a sequential version. We illustrate the gain (both in terms of complexity and performance) with respect to other existing methods.

The paper is organized as follows. In Section II, we introduce the system model and state the CA ML ambiguity-resolution problem. In Section III, we derive and discuss a CA ML ambiguity-resolution algorithm based on the minimization of the constrained Bethe free energy of the system. In Section IV, we focus on the sequential implementation of the proposed CA ML ambiguity-resolution method. Finally, in Section V we illustrate the performance of our approach by simulation results.

Notations: The notational conventions adopted in this paper are as follows. Italic indicates a scalar quantity, as in a or A; boldface lowercase indicates a vector quantity, as in $\mathbf{a}$; the kth element of vector $\mathbf{a}$ is denoted a_k; capital normal and boldface letters respectively indicate random variables and vectors, as in A and $\mathbf{A}$; calligraphic letters represents the set of values that a random variable or vector can take on: for example $\mathcal{A}$ is the set of possible values of $\mathbf{A}$; the estimate of a vector $\mathbf{a}$ is written $\hat{\mathbf{a}}$. $|\mathcal{A}|$ denotes the number of elements in $\mathcal{A}$. $\|\cdot\|$ is the ℓ_2-norm of a vector. The probability of a random vector $\mathbf{A}$ evaluated at $\mathbf{a}$ is denoted $p_{\mathbf{A}}(\mathbf{a})$. Finally, $\propto$ denotes equality up to a positive normalization factor.

II. System Model and Problem Statement

A. System Model

We consider a digital communication scheme in which a sequence of information bits, say $\mathbf{u}$, has to be transmitted through an AWGN channel. We assume that the sequence of information bits is protected against channel disturbances by a rate-R error-correcting code $\chi(\cdot)$, mapping the information sequence

onto a sequence of coded bits $\mathbf{x} = \chi(\mathbf{u})$. The coded bits are then mapped onto a constellation alphabet Ω, leading to a sequence of K complex symbols $\mathbf{a}$. The symbols are finally shaped by a unit-energy root-raised cosine pulse $c(t)$ with roll-off ζ, before being transmitted through the noisy channel. At the receiver side, the baseband received signal may be expressed as

$$\tilde{r}(t) = \sum_k a_k \, c(t - kT - \tau) \, e^{j\theta} + \tilde{w}(t), \tag{1}$$

where T is the symbol period, τ is the channel delay, θ is the carrier phase offset and $\tilde{w}(t)$ is the envelope of an additive white complex Gaussian noise with two-sided pass-band power spectral density $N_0/2$. The received signal is passed through a low-pass filter with cut-off frequency $f_c > (1+\zeta)/2T$ and is sampled at a rate $T_s^{-1} = 2f_c$:

$$r(lT_s) = \sum_k a_k \, c(lT_s - kT - \tau) e^{j\theta} + w(lT_s). \tag{2}$$

It can be shown (see *e.g.,* [33]) that the noise samples $w(lT_s)$ are independent complex-valued zero-mean Gaussian variables with variance $2N_0/T_s$. We define $\mathbf{r}$ as the vector stacking the observation samples $r(lT_s)$.

Since the sampling period satisfies the Nyquist-Shannon condition, $\mathbf{r}$ is a sufficient statistic of $r(t)$. Hence, we will consider the observation vector $\mathbf{r}$ instead of $r(t)$ in all our subsequent derivations. Note that this equivalence holds as long as l ranges from $-\infty$ to $+\infty$, *i.e.,* $\mathbf{r}$ contains an infinite number of samples. However, in practice, we can limit the size of $\mathbf{r}$ to a finite number of elements, say L, without any significant loss of precision.

In general, the carrier phase offset and the channel delay may be broken up as follows

$$\theta = k_\theta \, \Psi + \epsilon_\theta \qquad \text{with } -\tfrac{\Psi}{2} \leq \epsilon_\theta < \tfrac{\Psi}{2}, \tag{3}$$

$$\tau = k_\tau \, T + \epsilon_\tau \qquad \text{with } -\tfrac{T}{2} \leq \epsilon_\tau < \tfrac{T}{2}, \tag{4}$$

where k_τ and k_θ are integers, and Ψ is the smallest angle of symmetry of constellation Ω. For example, for phase shift keying (PSK) we have $\Psi = 2\pi/|\Omega|$ whereas for quadrature amplitude (QAM) constellation $\Psi = \pi/2$. We can give the following interpretation to k_τ and k_θ: k_τ is the integer number of symbol periods in the overall channel delay τ; k_θ represents the sector of the complex plane in which the carrier phase is located. In the sequel, we assume that the fractional timing and phase offset, *i.e.,* respectively ϵ_τ and ϵ_θ, can be properly estimated by means of conventional blind synchronizers [34], [35], so that

they can be neglected[2]. Under this assumption, the only remaining unknown synchronization parameters are k_τ and k_θ. These parameters are usually referred to as *timing* and *phase ambiguities*. For the sake of notational convenience, we will stack k_τ and k_θ in a vector $\mathbf{b}$, which can take values in a set $\mathcal{B}$. The focus of this paper is on determining $\mathbf{b}$ from $\mathbf{r}$.

B. Data Detection and Ambiguity Resolution

The main goal of the receiver is to recover the information bits $\mathbf{u}$. The observation vector $\mathbf{r}$ is usually processed in two steps: *i)* first, an estimate $\hat{\mathbf{b}}$ of the unknown synchronization parameters is computed; *ii)* a decision $\hat{\mathbf{u}}$ about the transmitted sequence is made through a decision rule, given the estimate $\hat{\mathbf{b}}$. For example, one can make a decision about the kth bit by maximizing the corresponding marginal a posteriori probability:

$$\hat{u}_k = \underset{u_k \in \{0,1\}}{\arg\max}\; p_{\mathrm{U}_k|\mathbf{R},\mathbf{B}}(u_k|\mathbf{r},\hat{\mathbf{b}}). \tag{5}$$

In the remainder of this paper, we will refer to (5) as conditional MAP (CMAP) bit-decision rule. Despite its suboptimality, CMAP has been shown to lead to outstanding performance, and CMAP-based receivers have a long tradition in practical receivers [4], [5], [36].

As far as CMAP detection is concerned, the quality of the final sequence decision depends on the accuracy of the decision made about $\hat{\mathbf{b}}$. In practice, a failure in properly estimating $\mathbf{b}$ leads to a wrong decision about $\mathbf{u}$ with probability almost equal to 1. In this context a desirable feature for an estimator of $\mathbf{b}$ is to minimize the probability of making a wrong decision about the parameter value. Given a decision rule $h(\cdot)$: $\hat{\mathbf{b}} = h(\mathbf{r})$, the associated error probability is given by

$$P_{\mathbf{e_b}|h(\cdot)} = 1 - \int_{\mathbb{R}^L} p_{\mathbf{B}|\mathbf{R}}(h(\mathbf{r})|\mathbf{r})\, p_{\mathbf{R}}(\mathbf{r})\, d\mathbf{r}. \tag{6}$$

From (6), it is clear that the decision rule that minimizes the probability of error is the MAP detector:

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}\in\mathcal{B}}\; p_{\mathbf{B}|\mathbf{R}}(\mathbf{b}|\mathbf{r}), \tag{7}$$

where $\mathcal{B}$ is the set of possible values of $\mathbf{b}$.

[2]Note that this assumption is not always satisfied in practice since the estimation of the blind synchronizers is not perfect and the synchronization parameters may be time-varying. The approach proposed in this paper can then be coupled with state-of-the-art synchronization techniques dealing with the estimation of the residual ϵ_θ, ϵ_τ. However, we do not consider this scenario hereafter for the sake of keeping the presentation as simple as possible.

If no a priori side information $p_{\mathbf{B}}(\mathbf{b})$ is available, the MAP detector reduces to the ML detector:

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}\in\mathcal{B}} p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}). \tag{8}$$

We will focus on ML detection, and note that the extension to MAP is straightforward. The ML criterion is the core of a number of ambiguity-resolution methods proposed in the literature, such as [20]–[22]. In these contributions, the ML solution is assumed to be intractable because the evaluation of the likelihood function $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ requires a summation over all possible sequences, *i.e.*,

$$p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) = \sum_{\mathbf{u}\in\mathcal{U}} p_{\mathbf{R},\mathbf{U}|\mathbf{B}}(\mathbf{r},\mathbf{u}|\mathbf{b}). \tag{9}$$

In this paper, we consider ML ambiguity resolution as a *free-energy* minimization problem and emphasize that this problem can be solved by polynomial-time algorithms. In particular, using the sum-product algorithm (SPA) we will show that, if the considered factor graph is cycle-free, the *exact* ML solution can be computed with a complexity lower than previously-proposed ambiguity-resolution methods. If the factor graph contains cycles, the proposed method provides an approximation of the ML estimate with a reasonable computational complexity.

III. Ambiguity Resolution Based on Free Energy Minimization

In this section, we show how ML ambiguity resolution methods can be implemented[3] within the factor-graph (FG) framework and its associated sum-product algorithm (SPA). First, we briefly recall the basics about FG representation and SPA message-update rules. This is followed by a short discussion, linking the SPA to free-energy minimization. Then, we describe how the considered ML problem may equivalently be regarded as a *free-energy* minimization problem and propose different expressions of the constrained Bethe free energy which only depend on a subset of belief normalization factors. Finally, in the last part of this section, we discuss the implementation and the complexity of the proposed algorithm.

A. Basics of Factor Graphs and the SPA

Let $v_1, v_2, \ldots, v_N$ denote a collection of variables and let $g(v_1, v_2, \ldots, v_N)$ denote a global distribution which may be factorized as

$$g(v_1, v_2, \ldots, v_N) = \frac{1}{Z}\prod_{j=1}^{M} f_j(v_{Q_j}), \tag{10}$$

[3]Exactly or approximately, depending on whether the considered factor graph has cycles or not.

where v_{Q_j} are sets of elements from $\{v_1, v_2, \ldots, v_N\}$ and $Z = \sum_{v_1,\ldots,v_N} g(v_1, v_2, \ldots, v_N)$ is a normalization factor. In many applications we are interested in efficiently computing the marginal functions of $g(v_1, v_2, \ldots, v_N)$:

$$g(v_i) = \sum_{\sim\{v_i\}} g(v_1, v_2, \ldots, v_N), \tag{11}$$

where the notation $\sim \{v_i\}$ denotes the summation over all the variables *except* v_i. The FG representation and the associated SPA [37], [38] provide a general framework to efficiently solve this problem: an FG is a bipartite graph that expresses the structure of the factorization (10). An FG has a variable vertex (or node) for each variable v_i, a factor node for each function f_j and an edge connecting variable node v_i to factor node f_j if and only if v_i is an argument of f_j. The SPA is an efficient procedure which enables to compute (either exactly or approximately) the marginals of the global function by passing messages along the edges of the corresponding factor graph. Denoting by $\mu_{v_i \to f_j}(v_i)$ the message sent from node v_i to node f_j and by $\mu_{f_j \to v_i}(v_i)$ the message sent from node f_j to node v_i, the message computations performed by the SPA may be expressed as follows:

$$\mu_{v_i \to f_j}(v_i) = \prod_{z \in n(v_i)\setminus\{j\}} \mu_{f_z \to v_i}(v_i), \tag{12}$$

$$\mu_{f_j \to v_i}(v_i) = \kappa_{ji}^{-1} \sum_{\sim\{v_i\}} \left(f_j(v_{Q_j}) \prod_{z \in n(f_j)\setminus\{i\}} \mu_{v_z \to f_j}(v_z) \right), \tag{13}$$

where $n(q)$ denotes the set of the neighbor indices of node q and κ_{ji} is an arbitrary positive constant.

When the FG is finite and cycle-free, the SPA can compute in a finite number of steps the exact marginals of the function that the graph represents [37], [38]. These marginals are equal, up to a normalization factor $\rho_i \triangleq \sum_{v_i} \prod_{h \in n(v_i)} \mu_{f_h \to v_i}(v_i)$, to the product of the messages entering each variable node:

$$g(v_i) = \rho_i^{-1} \prod_{h \in n(v_i)} \mu_{f_h \to v_i}(v_i). \tag{14}$$

It can also be shown that marginals of v_{Q_j} can be obtained as

$$g(v_{Q_j}) = \rho_j^{-1} f_j(v_{Q_j}) \prod_{i \in n(f_j)} \mu_{v_i \to f_j}(v_i), \tag{15}$$

where $\rho_j \triangleq \sum_{v_{Q_j}} f_j(v_{Q_j}) \prod_{i \in n(f_j)} \mu_{v_i \to f_j}(v_i)$ is a normalization factor.

When the FG has cycles, the SPA is not guaranteed to deliver the exact marginals, but only approximations thereof. We call the resulting solutions the *beliefs*[4]:

$$b_{Q_j}^{\mathrm{SPA}}(v_{Q_j}) = \rho_j^{-1}\, f_j(v_{Q_j}) \prod_{i \in n(f_j)} \mu_{v_i \to f_j}(v_i), \tag{16}$$

$$b_i^{\mathrm{SPA}}(v_i) = \rho_i^{-1} \prod_{j \in n(v_i)} \mu_{f_j \to v_i}(v_i). \tag{17}$$

Note that, in the presence of cycles in the FG, the SPA becomes an iterative algorithm. The question of its fixed points and its convergence will be discussed in the next section.

B. Free Energy Minimization

A new interpretation of the SPA was recently offered in [25]. For a factorization as (10), we introduce the Bethe free energy as

$$\begin{aligned} F_{\mathrm{Bethe}}(\{b_{Q_j}\}_j, \{b_i\}_i) = &- \sum_{j=1}^{M} \sum_{v_{Q_j}} b_{Q_j}(v_{Q_j}) \log f_j(v_{Q_j}) \\ &+ \sum_{j=1}^{M} \sum_{v_{Q_j}} b_{Q_j}(v_{Q_j}) \log b_{Q_j}(v_{Q_j}) \\ &- \sum_{i=1}^{N} (d_i - 1) \sum_{v_i} b_i(v_i) \log b_i(v_i), \end{aligned} \tag{18}$$

where d_i denotes the degree of (*i.e.,* the number of edges connected to) variable node v_i. Now, suppose we try to minimize the Bethe free energy with respect to the functions $\{b_{Q_j}\}_j, \{b_i\}_i$, subject to the following normalization and consistency constraints

$$\sum_{v_{Q_j}} b_{Q_j}(v_{Q_j}) = 1 \qquad \forall\, j, \tag{19}$$

$$\sum_{v_i} b(v_i) = 1 \qquad \forall\, i, \tag{20}$$

$$\sum_{\sim\{v_i\}} b_{Q_j}(v_{Q_j}) = b_i(v_i) \qquad \forall\, j, \forall\, i \in Q_j, \forall\, v_i. \tag{21}$$

It was then shown in [25] that the SPA, provided it converges, leads to beliefs (16)-(17) which are *stationary points* of the constrained Bethe free energy. It was also shown that when the factor graph is

[4]The notation used for the beliefs in (16)-(17) should not be confused with the vector of synchronization parameters $\mathbf{b}$.

cycle-free, the minimized constrained Bethe free energy can be related to Z through:[5]

$$F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}}\right\}_j, \{b_i^{\text{SPA}}\}_i) = -\log Z. \tag{22}$$

When the FG contains cycles, the relation (22) no longer holds. However, $F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}}\right\}_j, \{b_i^{\text{SPA}}\}_i)$ can be interpreted as an *approximation* of $-\log Z$, just as the beliefs are interpreted as approximations of the exact marginals. We will use this result in the next sections to derive new phase and timing ambiguity resolution methods.

In [25], the authors also emphasize that the SPA is not ensured to converge for arbitrary FGs. In particular, the beliefs computed by the SPA are neither ensured to satisfy (21) nor to monotonically decrease the Bethe free energy throughout the iterations. Following this result, several variants of the SPA, which are guaranteed to converge, have been proposed in the literature, see *e.g.*, [39]–[41]. The results which will be derived in the rest of this paper also apply to these algorithms because their fixed points are stationary points of the constrained Bethe free energy. However, for the sake of keeping the discussion as simple as possible, we will only refer the SPA hereafter.

Before concluding this section, let us mention that the choice of the κ_{ji}'s in (12) does not affect the fixed points of the SPA [25]. However, a proper choice of the κ_{ji}'s can greatly influence the convergence of the SPA in cyclic FGs. We will see in the sequel that these factors also play an important role in the efficient evaluation of the minimum Bethe free energy associated to the FG.

C. Ambiguity Resolution through Free-energy Minimization

In this section, we demonstrate how ML ambiguity resolution can be related to free-energy minimization. Consider the joint distribution $p_{\mathbf{U}|\mathbf{R},\mathbf{B}}(\mathbf{u}|\mathbf{r},\mathbf{b})$, the marginals of which are exactly those used in the CMAP detector (5). We can rewrite this distribution as

$$p_{\mathbf{U}|\mathbf{R},\mathbf{B}}(\mathbf{u}|\mathbf{r},\mathbf{b}) = \frac{1}{p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})} p_{\mathbf{U},\mathbf{R}|\mathbf{B}}(\mathbf{u},\mathbf{r}|\mathbf{b}). \tag{23}$$

Assuming we can factorize $p_{\mathbf{U},\mathbf{R}|\mathbf{B}}(\mathbf{u},\mathbf{r}|\mathbf{b})$, we can make the following association between (10) and (23):

$$\begin{aligned} p_{\mathbf{U}|\mathbf{R},\mathbf{B}}(\mathbf{u}|\mathbf{r},\mathbf{b}) &\leftrightarrow g(v_1, v_2, \ldots, v_N) \\ \frac{1}{p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})} &\leftrightarrow \frac{1}{Z} \\ p_{\mathbf{U},\mathbf{R}|\mathbf{B}}(\mathbf{u},\mathbf{r}|\mathbf{b}) &\leftrightarrow \prod_j f_j(v_{Q_j}). \end{aligned}$$

[5]In this case, the Bethe free energy is also known as the Helmholtz-Gibbs free energy.

Taking (22) into account, the ML ambiguity resolution problem can therefore be rewritten as:

$$\begin{aligned}\hat{\mathbf{b}} &= \arg\max_{\mathbf{b}\in\mathcal{B}} \log p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}),\\ &= \arg\max_{\mathbf{b}\in\mathcal{B}} \log Z,\\ &= \arg\min_{\mathbf{b}\in\mathcal{B}} F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\text{SPA}|\mathbf{b}}\right\}_i).\end{aligned}$$

The last transition is exact when the FG is cycle-free and approximate otherwise (see section III-B). $b_{Q_j}^{\text{SPA}|\mathbf{b}}(u_{Q_j})$ and $b_i^{\text{SPA}|\mathbf{b}}(u_i)$ denote the beliefs obtained by the SPA after *convergence*[6] when applied to the FG of $p_{\mathbf{U},\mathbf{R}|\mathbf{B}}(\mathbf{u},\mathbf{r}|\mathbf{b})$. The notation $(\cdot|\mathbf{b})$ indicates that the beliefs are conditioned on $\mathbf{b}$.

In the sequel, with a slight abuse of language we will refer to $\hat{\mathbf{b}}$ as the ML estimate in both the exact and approximate cases. In practice this leads to the following technique to determine $\mathbf{b}$: *i)* for every possible value of $\mathbf{b}$, determine the minimal constrained Bethe free energy $F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\text{SPA}|\mathbf{b}}\right\}_i)$; *ii)* the ML estimate of $\mathbf{b}$ is the value which gives rise to the smallest minimal constrained Bethe free energy.

D. Alternative Expressions of the Bethe Free Energy

In the previous section, we emphasized the relation between the ML ambiguity resolution problem and the minimization of the Bethe free energy with respect to $\mathbf{b}$. Unfortunately, a direct evaluation of the Bethe free energy via (18) can often be cumbersome in terms of storage and computation. In this section, we propose alternative expressions for the evaluation of $F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\text{SPA}|\mathbf{b}}\right\}_i)$ which only require the evaluation of a subset of normalization factors ρ_i, ρ_j. These expressions are based on the following results:

Proposition 3.1: Let $\rho_j(\mathbf{b})$ and $\rho_i(\mathbf{b})$ be the normalization factors associated to beliefs $b_{Q_j}^{\text{SPA}|\mathbf{b}}$ and $b_i^{\text{SPA}|\mathbf{b}}$. Then,

$$\begin{aligned}&F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\text{SPA}|\mathbf{b}}\right\}_i) =\\ &-\sum_{j=1}^{M}\log\rho_j(\mathbf{b}) + \sum_{i=1}^{N}(d_i-1)\log\rho_i(\mathbf{b}).\end{aligned} \tag{24}$$

□

[6]As mentioned in Section III-B, a stationary point of the Bethe free energy is achieved only if the SPA is at a fixed point. In practice, convergence is often assumed when the variation of the (normalized) SPA messages drops below some threshold.

Proposition 3.2: If $j \in n(v_i)$, then the following equality holds:

$$\frac{\rho_j(\mathbf{b})}{\rho_i(\mathbf{b})} = \kappa_{ji}. \tag{25}$$

□

The proofs can be found in Appendices A and B. These propositions can somehow be regarded as corollaries of the results proved in [25]. Proposition 3.1 provides an expression of the constrained Bethe free energy which only depends on normalization factors $\rho_j(\mathbf{b})$, $\rho_i(\mathbf{b})$. Proposition 3.2 gives a connection between the normalization factors of adjacent nodes. Note, on the one hand, that the relation between these factors only depends on the constants κ_{ji}'s appearing in (12). On the other hand, these constants can be set arbitrarily without affecting the fixed points of the SPA[7]. It is therefore tempting to try to simplify (24) by assigning "well-chosen" values to the κ_{ji}'s. We give hereafter two examples of such simplifications:

- *Constant κ_{ji}'s:* $\kappa_{ji} = \kappa$ is a constant $\forall i, j$. If we set $\kappa = 1$, we have from (25) that all the normalization factors are equal:

  $$\rho_i(\mathbf{b}) = \rho_j(\mathbf{b}) \triangleq \rho(\mathbf{b}) \qquad \forall i, j. \tag{26}$$

  This implies that

  $$-F_{\mathrm{Bethe}}(\left\{b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\mathrm{SPA}|\mathbf{b}}\right\}_i) = (M + N - \sum_{i=1}^{N} d_i) \log \rho(\mathbf{b}). \tag{27}$$

  The Bethe free energy can therefore be evaluated by computing *one* normalization factor (instead of $M + N$). Note that $M + N = 1 + \sum_i d_i$ for any acyclic graph, and therefore (27) reduces to

  $$-F_{\mathrm{Bethe}}(\left\{b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\mathrm{SPA}|\mathbf{b}}\right\}_i) = \log \rho(\mathbf{b}). \tag{28}$$

- *Constant κ_{ji}'s on subtrees*: we focus on the case of the FG comprising two components G_1 and G_2, both of which are trees[8]. The graphs G_1 and G_2 are connected with one another through N_{c}

[7]Some care has however to be taken to ensure the convergence of the SPA. One can for example take a suitable cut of the FG and normalize the messages crossing this cut to some value. See, for example, case 2 in Section III-E.

[8]The reasoning hereafter can be easily extended to FGs made up of more than 2 trees.

(common) variable nodes, say $v_1, \ldots, v_{N_c}$. We can perform the SPA using

$$\kappa_{ji} = \begin{cases} \sum_{v_{Q_j}} f_j(v_{Q_j}) \prod_{z \in n(f_j)\setminus\{i\}} \mu_{v_z \to f_j}(v_z) & \forall i \in \{1, \ldots, N_c\},\ \forall j \in n(v_i) \\ 1, & \text{otherwise.} \end{cases}$$

This choice of κ_{ji}'s is equivalent to: *i)* normalizing to 1 the messages entering nodes $v_1, \ldots, v_{N_c}$, *i.e.,*

$$\sum_{v_i} \mu_{f_j \to v_i}(v_i) = 1, \qquad \forall i \in \{1, \ldots, N_c\},\ \forall j \in n(v_i), \tag{29}$$

ii) leaving all other messages unnormalized ($\kappa_{ji} = 1$). From (25), all the normalization factors in G_1 and G_2 are constant. Hence, there are $N_c + 2$ distinct normalization factors: $\rho_{G_1}(\mathbf{b})$ for nodes in G_1, $\rho_{G_2}(\mathbf{b})$ for nodes in G_2 and $\rho_i(\mathbf{b})$ for v_i's with $i \in \{1, \ldots, N_c\}$. Particularizing (24), we find that

$$-F_{\text{Bethe}}(\left\{b_{Q_j}^{\text{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\text{SPA}|\mathbf{b}}\right\}_i) = \log \rho_{G_1}(\mathbf{b}) + \log \rho_{G_2}(\mathbf{b}) - \sum_{i=1}^{N_c} (d_i - 1) \log \rho_i(\mathbf{b}), \tag{30}$$

which is computationally much more efficient than (24) for general FGs.

E. Implementation and Complexity

In the previous section, we emphasized that the minimum constrained Bethe free energy can be evaluated from a subset of the belief normalization factors, see *e.g*, (27) and (30). This results can be applied to the problem of phase and timing ambiguity resolution. We distinguish between three cases:

1) *Acylic FGs*: this case corresponds, for example, to BPSK transmissions with convolutional error correcting code. Since the FG is acyclic, the minimum Bethe free energy is equal to $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$. From (28), it can be evaluated from the knowledge of *one* (arbitrary) normalization factor $\rho(\mathbf{b})$ if we set $\kappa_{ji} = 1$ for all nodes in the FG. Note that the computation of one single normalization factor only requires to evaluate the SPA messages in *one* direction in the FG.
2) *Cyclic FGs with Acyclic Subgraphs*: this is the case of turbo decoders, where the FG of the two constituent BCJR decoders are cycle-free and connected by N_c information-bit nodes. It is very tempting to apply (28) since it only requires the evaluation of one normalization factor. However, (28) is based on the hypothesis that the messages are not renormalized through the SPA iterations (*i.e.,* $\kappa_{ji} = 1\ \forall i, j$). If the FG is cyclic, this approach usually fails for stability reasons: the SPA messages converge to 0 or ∞. Instead, the minimum Bethe free energy can be evaluated through

(30), which implies a renormalization of the messages leaving the N_c connecting nodes. This renormalization prevents the SPA messages from diverging to 0 or ∞ and the stability is therefore ensured.

3) *General FGs*: when the FG does not have any particular structure, the Bethe free energy can be evaluated by means of (24). This is for example the case of transmissions using LDPC codes.

It is interesting to relate the complexity of the proposed ambiguity resolution method to the complexity of the CMAP receiver (5). First, note that $b_i^{\mathrm{SPA}|\mathbf{b}}(u_i)$ corresponds (exactly or approximately) to the conditional a posteriori probability $p_{\mathrm{U}_k|\mathbf{R},\mathbf{B}}(u_k|\mathbf{r},\mathbf{b})$ considered in the CMAP receiver (5). With a slight abuse of language, we will therefore associate the complexity of the (possibly approximate) CMAP receiver to the task of computing beliefs $b_i^{\mathrm{SPA}|\mathbf{b}}(u_i)\ \forall u_i \in \{0,1\}, \forall i$.

As mentioned in Section III-C, the beliefs computed by the SPA are also those minimizing the constrained Bethe free energy. Therefore, the complexity of the proposed ambiguity-resolution methods is at most equal to $|\mathcal{B}|$ times the complexity of the CMAP receiver.

In some cases (see Section III-D), the Bethe free energy can be evaluated from a subset of normalization factors and the complexity can then be reduced. For example, in the cycle-free case only one normalization factor is required to evaluate the Bethe free energy. We can thus save 50% of the computations[9] since the messages on each edge have only to be computed in one direction.

As a point of comparison, we can note that most CA ambiguity-resolution methods proposed so far in the literature require one decoding operation per possible value of $\mathbf{b}$. In particular, considering methods applying to CMAP-based receivers (5), the ML-based approaches proposed in the literature, see *e.g.*, [19], [21], exhibit a complexity equal to $|\mathcal{B}|$ times the complexity of one CMAP decoding operation. From our previous reasoning, we come therefore to the conclusion that the proposed method has a computational complexity equal or even lower (in the cycle-free case) than these methods.

Finally, we mention that the proposed ambiguity resolution method can be implemented from the matched-filter outputs $\mathbf{y}(\mathbf{b})$, where

$$y_k(\mathbf{b}) \triangleq T_s \sum_l r(lT_s)\, c(lT_s - kT - k_\tau T)\, e^{-jk_\theta \Psi}. \tag{31}$$

[9]More precisely, the complexity of the ambiguity-resolution method is $\frac{|\mathcal{B}|}{2}$ times the complexity of the CMAP-receiver since the evaluation of *one* normalization factor only requires to compute the SPA messages in one direction in the FG. However, once the decision about $\mathbf{b}$ has been made, the evaluation of the beliefs $b_i^{\mathrm{SPA}|\mathbf{b}}(u_i)$ requires to compute the SPA messages in the other direction for $\mathbf{b} = \hat{\mathbf{b}}$. The complexity of the overall receiver is therefore equal to $\frac{|\mathcal{B}|+1}{2}$ times the complexity of the CMAP-receiver.

More particularly, we show in appendix C that

$$p_{\mathbf{R},\mathbf{U}|\mathbf{B}}(\mathbf{r},\mathbf{u}|\mathbf{b}) \propto \gamma(\mathbf{b})\, p_{\mathbf{Y},\mathbf{U}|\mathbf{B}}(\mathbf{y}(\mathbf{b}),\mathbf{u}|\mathbf{b}), \tag{32}$$

where

$$\gamma(\mathbf{b}) = \exp\left(\frac{\|\mathbf{y}(\mathbf{b})\|^2}{2N_0}\right). \tag{33}$$

Working with the matched-filter outputs $\mathbf{y}(\mathbf{b})$ instead of $\mathbf{r}$ is usually more convenient for finding a nice factorization of the objective function (10): *i)* the length of $\mathbf{y}$ is T_s/T times the length of $\mathbf{r}$, which is interesting since $T_s \leq T$; *ii)* unlike $r(lT_s)$, $y_k(\mathbf{b})$ only depends on *one* single data symbol.

IV. Sequential ML Ambiguity Resolution

In the previous section, we proposed an SPA framework for the resolution of phase/timing ambiguity problems by Bethe free-energy minimization. In this section, we propose a low-complexity version of this method by computing the normalization factors from a *subset* of the messages entering each node. This approach is shown to have a nice sequential implementation when the FG has a chain structure.

This section is organized as follows. We first expose the sequential computation of approximate likelihood functions in the case where the FG has a chain structure. Then, we elaborate on the choice of a relevant stopping criterion. Finally, we briefly discuss the application of the proposed sequential procedure to some cyclic FGs.

A. Sequential Approximated ML Solution

Consider a cycle-free FG made up of N (cycle-free) sub-FGs. Assume moreover that the FG has a chain structure as represented in Fig. 1. From our derivations in Section III we know that, if $\kappa_{ji} = 1$ $\forall i,j$, $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ is equal to the normalization factor of any belief in the FG, *i.e.*,

$$p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) = \rho_i(\mathbf{b}) \qquad \forall i \in \{1,\ldots,N\}, \tag{34}$$

where

$$\rho_i(\mathbf{b}) = \sum_{v_i} \prod_{z \in n(v_i)} \mu_{f_z \to v_i}(v_i). \tag{35}$$

At node v_i, we consider the following pseudo likelihood function:

$$p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) = \sum_{v_i} \prod_{z \in n^-(v_i)} \mu_{f_z \to v_i}(v_i). \tag{36}$$

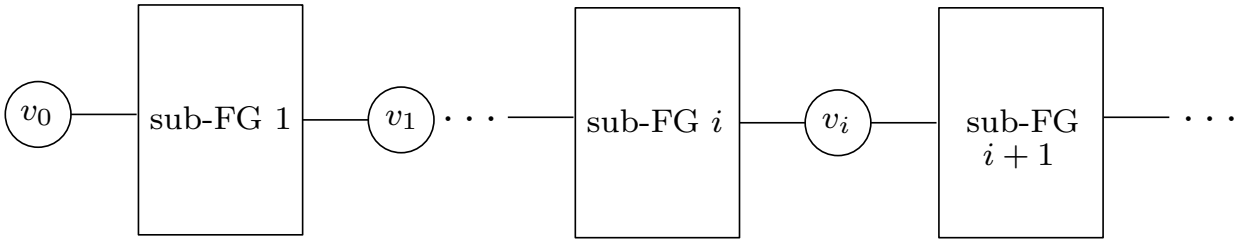

Figure 1. FG well-suited to a sequential implementation of the ML criterion.

where $n^-(v_i)$ represents the set of the neighbors of v_i which are in sub-FG i. Then, $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ can be regarded as an approximation of $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ where all the messages coming from the sub-FG $i+1$ are set to 1. The evaluation of $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ is less complex than the one of $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ since only the messages coming from sub-FG i need to be computed for the former. It is therefore very tempting to consider the following approximate ML problem to reduce the complexity:

$$\hat{\mathbf{b}}^{(i)} = \arg\max_{\mathbf{b}} p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}). \tag{37}$$

In practice, the choice of a proper approximation $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ is a tradeoff between accuracy and complexity. On the one hand, it is clear that the complexity associated to the evaluation of $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ increases with i. On the other hand, the quality of the approximation $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) \simeq p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ also improves with i: if $i = N$ we have $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) = p^{(N)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ since v_N is the last node in the FG; on the contrary we have $p^{(0)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b}) = |\mathcal{B}|^{-1}$ and is therefore a very poor approximation of $p_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$.

It is important to note that the evaluation of $p^{(i+1)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ can be made with a limited number of operations from the knowledge of $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$. Indeed, messages $\mu_{f_z \to v_{i+1}}(v_{i+1})$, $z \in n^-(v_{i+1})$, can be computed from messages $\mu_{f_z \to v_i}(v_i)$, $z \in n^-(v_i)$ by applying the SPA on sub-FG $i+1$. In other words, all the operations made to evaluate $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ can be reused in the evaluation of $p^{(i+1)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$. We propose therefore to following sequential procedure:

1) Initialize $i = 0$.
2) Solve (37).
3) If $\hat{\mathbf{b}}^{(i)}$ satisfies a stopping criterion, stop the computation; otherwise, set $i = i + 1$ and go to step 2.

The choice of the stopping criterion is discussed in the next section.

B. The Stopping Criterion

In practice, we wish to stop the computation when the probability of making a wrong decision is "low enough". In other words, the confidence about the ambiguity should be large enough to enable to recover (if possible) the performance of the perfectly synchronized system. Denoting $\mathrm{FER}_{\mathrm{perf}}$ the frame-error rate achieved by the perfectly-synchronized system, this means, we want to find a decision rule $h(\cdot) : \hat{\mathbf{b}} = h(\mathbf{r})$ such that

$$P_{\mathrm{e_b}|h(\cdot)} \ll \mathrm{FER}_{\mathrm{perf}}, \tag{38}$$

where $P_{\mathrm{e_b}|h(\cdot)}$ has been defined in (6). The idea behind (38) is as follows: if the probability of making a wrong decision on $\mathbf{b}$ is much smaller than $\mathrm{FER}_{\mathrm{perf}}$, the synchronization operation will not significantly degrades the performance achievable by the perfectly-synchronized system.

Condition (38) is satisfied if we find a decision rule $h(\mathbf{r})$ such that

$$1 - p_{\mathbf{B}|\mathbf{R}}(h(\mathbf{r})|\mathbf{r}) \ll \mathrm{FER}_{\mathrm{perf}} \qquad \forall\, \mathbf{r}. \tag{39}$$

Therefore, a good stopping criterion for our sequential implementation would consist in stopping the computation as soon as

$$1 - p_{\mathbf{B}|\mathbf{R}}(\hat{\mathbf{b}}^{(i)}|\mathbf{r}) \ll \mathrm{FER}_{\mathrm{perf}}. \tag{40}$$

This way, we are ensured to have a quality of decision not affecting the achievable frame-error-rate. Unfortunately, (40) requires to compute probability $p_{\mathbf{B}|\mathbf{R}}(\mathbf{b}|\mathbf{r})$ and is therefore not a very useful stopping criterion in a sequential implementation. Instead, since a new estimate is computed at each step by maximizing $p^{(i)}_{\mathbf{B}|\mathbf{R}}(\mathbf{b}|\mathbf{r})$ (or $p^{(i)}_{\mathbf{R}|\mathbf{B}}(\mathbf{r}|\mathbf{b})$ if the prior is uniform), we propose the following stopping criterion:

$$\begin{aligned} 1 - p^{(i)}_{\mathbf{B}|\mathbf{R}}(\hat{\mathbf{b}}^{(i)}|\mathbf{r}) &= 1 - \max_{\mathbf{b}} p^{(i)}_{\mathbf{B}|\mathbf{R}}(\mathbf{b}|\mathbf{r}) \\ &\ll \mathrm{FER}_{\mathrm{perf}}. \end{aligned} \tag{41}$$

Of course, (41) does not necessarily imply (40). However, the validity of this stopping criterion will be assessed by simulations in Section V.

C. Sequential Ambiguity Resolution on Cyclic FGs

The sequential ML approach appears quite appealing since it can dramatically reduce the complexity of the code-aided ambiguity-resolution method. As mentioned in Section IV-A, its implementation however requires that the considered FG has a chain structure. This is for example the case for convolutionally-coded transmissions, see *e.g.,* [37]. On the contrary, the FGs associated to turbo-coded or LDPC-coded

transmissions contain a lot of cycles. In such cases, the implementation of the sequential algorithm is in principle not possible due to the non-sequential structure of the FG.

In some situations, it is however possible to recover a sequential structure by making some additional approximations. For example, in the case of a turbo-coded transmission, we can decide to apply the sequential algorithm on one of the two constituent decoders. Of course, taking this approach, we are no longer solving the initial free-energy minimization problem but rather an approximation of it. However, considering this approximated problem may turn out to be sufficient to solve the ambiguity problem and has the advantage to dramatically reduce the computational complexity. We will illustrate this approach by simulation in the next section.

V. Simulation Results

In this section, the BER performance of the data-aided (DA) [1]–[3], the expectation-maximization (EM) [21], and the free energy minimization (FE) algorithms are evaluated and compared to that of the perfect synchronization, *i.e.,* to the case where the decoder knows perfectly timing and phase ambiguities. We consider the cases of convolutionally-coded and turbo-coded transmissions.

A. Convolutional Codes

We first study the performance of the proposed ambiguity-resolution algorithm in the case of a convolutionally-coded BPSK transmission. The FG corresponding to this type of code is cycle-free and periodic. The FE algorithm implements therefore the exact MAP criterion. Moreover, the periodicity of the FG allows for the implementation of the sequential approach described in section IV.

We consider a rate-$\frac{1}{3}$ systematic convolutional code with encoding polynomial (21,37). The length of the coded sequence is set to 128. We use (28) to evaluate the a posteriori probability of the synchronization parameters, *i.e.,* only one normalization factor is needed. The timing and phase ambiguity resolutions are treated separately. We consider three possible timing ambiguities (namely $k_\tau \in \{0, 1, 2\}$). Since BPSK modulation is used, phase ambiguity can take on two values, *i.e.,* $k_\theta \in \{0, 1\}$.

Fig. 2 represents the probability of wrong synchronization achieved by the EM and FE algorithms. We can notice that FE exhibits the best performance. Indeed, since the considered system has a cycle-free FG representation, the proposed Bethe-free energy algorithm reduces to MAP estimation which is optimal in terms of minimization of the probability of wrong synchronization. Note moreover that the complexity of the FE algorithm is half the one of the EM algorithm (see section III-E) since the a posteriori probability of the synchronization parameters can be efficiently evaluated via (28).

We can also notice a significant difference between the performance achieved for PAR and TAR. Since we are implementing the optimal MAP receiver, the degradation of the performance for TAR is only due to the code structure. In other words, given the considered code, no ambiguity-resolution algorithms can perform better than the proposed FE algorithm. It is worth noticing that some codes are badly suited to the resolution of ambiguities. For example, a linear code which contains the "all-one" word in its codebook can never resolve phase ambiguities for a BPSK transmission: flipping all the bits of a codeword still leads to a valid codeword. In the same way, a rate-$\frac{1}{n}$ convolutional code can never resolve timing ambiguities which are multiple of nT. The latter problem can be circumvent by interleaving the coded bits. Indeed, the presence of the interleaver breaks the "periodicity" of the code and allows therefore for timing ambiguity resolution. The curves labeled "TAR & interleaver" in Fig. 2 illustrates this effect: one can observe that the TAR can be properly resolved when using an interleaver at the output of the coder.

Fig. 3 represents the BER achieved by the system synchronized by the EM and FE algorithms. In the case of PAR, we note that the quality of synchronization is sufficiently high to recover the same BER as that of the perfectly-synchronized system. For TAR, the recovery of the BER of the perfectly-synchronized system requires the use of an interleaver at the output of the coder. The BER achieved by the system synchronized by the sequential FE algorithm described in Section IV is also represented. We note that the sequential approach does not lead to any significant degradation of the BER with respect to the standard FE algorithm.

The computational savings allowed by the sequential approach is illustrated in Fig. 4. The curves represent the cumulative distribution function (CDF) of the decoding stage i at which the sequential FE algorithm makes its final decision. Three different values of E_S/N_0 are considered. We see that the number of decoding stages decreases when the SNR increases. In the case of PAR, the sequential approach has to run until the end of the trellis for most of the realizations when $E_S/N_0 = -6$dB. On the other hand, the number of decoding stages never exceed 20 when $E_S/N_0 = 4$dB. We note similar results in the case of TAR (with interleaving).

B. Turbo Codes

We consider a rate-$\frac{1}{3}$ BPSK turbo code with encoding polynomial (21,37) for the constituent convolutional codes. Timing and phase ambiguity resolutions are treated separately. The possible timing ambiguities (resp. phase ambiguities) include 0, 1 and 2 (resp. 0 and 1). The decoder stops decoding after 5 turbo iterations. The timing and phase ambiguities are estimated by using the expression of the Bethe free energy stated in (30). We consider the case of data words of length 128 and 512.

Fig. 5, 6 and 7 represent the performance achieved by different ambiguity resolution methods (DA, EM, FE) for codewords of length equal to 128. In Fig. 5, the data-aided algorithm is evaluated at a pilot length of 10, 30 and 60. For the timing ambiguity resolution, the performance is obviously better when the pilot length increases. For the phase ambiguity resolution, increasing the pilot length from 30 to 60 yields worse performance because the E_S/N_0 compensation for the longer pilot sequence outweighs the better accuracy. It can be observed that timing ambiguity resolution requires a longer pilot length than phase ambiguity resolution to obtain performance close to the perfect synchronization.

In Fig. 6, the EM algorithm is evaluated by making a decision at decoder 1 or at decoder 2 after one turbo iteration. The EM algorithm performs close to the perfect synchronization for both timing and phase ambiguity resolutions when the decision is made at decoder 2. There is a performance loss of 1 dB when the decision is made at decoder 1 for phase ambiguity resolution. For TAR, making a decision at decoder 1 yields poor performance. This can be explained by the periodicity of the code trellis: shifting by n bits the output of a rate-$\frac{1}{n}$ convolutional encoder still leads to a valid codeword. Therefore, in the limit of an infinite sequence[10], timing ambiguities multiple of nT cannot be resolved for rate-$\frac{1}{n}$ convolutional codes by *any* ambiguity resolution method. The presence of the interleaver between the two constituent decoders breaks the code symmetry. This explains the improvement of the performance when the decision is made at the second decoder.

In Fig. 7, the Bethe free-energy minimization algorithm (see Section IV) is evaluated by making a decision at decoder 1 or at decoder 2 after one turbo iteration. The decision at decoder 1 only considers the SPA messages in the FG of the first convolutional code and computes the corresponding Bethe free energy with (28). The decision at decoder 2 is based on (30) and therefore exploits the messages from the two convolutional decoders. A sequential version of the algorithm is implemented as follows. For PAR, the sequential algorithm described in section IV is applied to the first convolutional decoder only. In the case of TAR, the sequential procedure is applied to the second convolutional decoder by taking the extrinsic probabilities computed by the first one into account.

We can make the following observations. Regarding PAR, the performance difference between the phase ambiguity resolution at decoder 1 and at decoder 2 is not significant. The performance achieved by the proposed approach is similar to that of the EM approach. On the other hand, TAR at the first decoder suffers from the same problem as EM: it cannot properly recover the timing ambiguity by exploiting the code structure of the convolutional code. When decision at decoder 2 is considered, we observe an

[10]The borders of the trellis break the periodicity and allows therefore for the resolution of some part of the ambiguity.

improvement of the BER with respect to EM. In particular, unlike EM, the proposed approach almost recovers the performance of the perfectly synchronized system for $E_S/N_0 \geq -1$dB.

We can also note the good behavior of the sequential approaches: the degradation with respect to the non-sequential procedure remains limited while the computational complexity of the ambiguity-resolution method decreases. In order to quantify this saving, Fig. 8 represents the CDF of the decoding stage at which the sequential FE algorithm makes its final decision. The synchronization decision is done obviously earlier as SNR increases. In particular, considering the case $E_S/N_0 = 0$dB for PAR (resp. TAR) we see that the final decision never exceeds 30 (resp. 60) trellis transitions.

Fig. 9 and 10 show respectively the BER performance achieved by EM and the proposed method when the length of the coded sequence is equal to 512. We observe that increasing the length of the codeword improves the effectiveness of the ambiguity resolution methods. The FE algorithm can recover the performance of the perfectly synchronized system. The EM algorithm also improves the performance but exhibits a slight degradation with respect to the perfectly-synchronized system at intermediate SNR for PAR. The sequential approach only leads to a negligible degradation.

VI. Conclusions

We have proposed a new class of code-aided ambiguity resolution algorithms, based on the connection between ML estimation, factor graphs, and free-energy minimization. This new class of algorithms can achieve good performance at a reasonable complexity cost, without relying on training sequences. We have also put forth a number of variations of these algorithms that are able to exploit the special structure in the underlying factor graph to reduce the computational complexity.

Appendix A

In this appendix, we give a proof of (24). Plugging the expressions (16)-(17) of $b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}(v_{Q_j})$ and $b_i^{\mathrm{SPA}|\mathbf{b}}(v_i)$ into the definition of the Bethe free energy (18), we obtain

$$\begin{aligned}
-F_{\mathrm{Bethe}}(\left\{b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}\right\}_j, \left\{b_i^{\mathrm{SPA}|\mathbf{b}}\right\}_i) = \\
&\sum_{j=1}^{M} \log \rho_j(\mathbf{b}) - \sum_{i=1}^{N} (d_i - 1) \log \rho_i(\mathbf{b}) \\
&- \sum_{j=1}^{M} \sum_{v_{Q_j}} b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}(v_{Q_j}) \sum_{i \in n(f_j)} \log \mu_{v_i \to f_j}(v_i) \\
&+ \sum_{i=1}^{N} (d_i - 1) \sum_{v_i} b_i^{\mathrm{SPA}|\mathbf{b}}(v_i) \sum_{j \in n(v_i)} \log \mu_{f_j \to v_i}(v_i). \qquad (42)
\end{aligned}$$

Let us show that the last two terms cancel out. First note that $\sum_{\sim\{v_i\}} b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}(v_{Q_j}) = b_i^{\mathrm{SPA}|\mathbf{b}}(v_i)$ from (21). Therefore,

$$\begin{aligned}&\sum_{j=1}^{M}\sum_{v_{Q_j}} b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}(v_{Q_j}) \sum_{i\in n(f_j)} \log \mu_{v_i\to f_j}(v_i)\\ &= \sum_{j=1}^{M}\sum_{i\in n(f_j)}\sum_{v_i} b_i^{\mathrm{SPA}|\mathbf{b}}(v_i)\, \log \mu_{v_i\to f_j}(v_i). \end{aligned} \tag{43}$$

In addition, we have that $\sum_{j=1}^{M}\sum_{i\in n(f_j)}$ is equal to $\sum_{i=1}^{N}\sum_{j\in n(v_i)}$ since both summations are equivalent to counting all the edges of the FG. Finally, taking (12) into account, we have

$$\begin{aligned}&\sum_{j=1}^{M}\sum_{v_{Q_j}} b_{Q_j}^{\mathrm{SPA}|\mathbf{b}}(v_{Q_j}) \sum_{i\in n(f_j)} \log \mu_{v_i\to f_j}(v_i)\\ &= \sum_{i=1}^{N}\sum_{j\in n(v_i)} (d_i-1)\sum_{v_i} b_i^{\mathrm{SPA}|\mathbf{b}}(v_i) \log \mu_{f_j\to v_i}(v_i), \end{aligned} \tag{44}$$

and therefore, the last two terms in (42) cancel out.

APPENDIX B

In this appendix, we give a proof of (25). Assume that (12)-(13) holds at every node in the FG. Using (16), we have for any factor node f_j:

$$\begin{aligned}\rho_j(\mathbf{b}) &= \sum_{v_{Q_j}} f_j(v_{Q_j}) \prod_{z\in n(f_j)} \mu_{v_z\to f_j}(v_z)\\ &= \sum_{v_i} \mu_{v_i\to f_j}(v_i) \sum_{\sim\{v_i\}} f_j(v_{Q_j}) \prod_{z\in n(f_j)\setminus\{i\}} \mu_{v_z\to f_j}(v_z)\\ &= \kappa_{ji} \sum_{v_i} \mu_{v_i\to f_j}(v_i)\, \mu_{f_j\to v_i}(v_i), \end{aligned} \tag{45}$$

where the last equality follows from (13). On the other hand, from (17) we have for any node v_i and any $j \in n(v_i)$:

$$\begin{aligned}\rho_i(\mathbf{b}) &= \sum_{v_i} \prod_{z\in n(v_i)} \mu_{f_z\to v_i}(v_i)\\ &= \sum_{v_i} \mu_{f_j\to v_i}(v_i) \prod_{z\in n(v_i)\setminus\{j\}} \mu_{f_z\to v_i}(v_i)\\ &= \sum_{v_i} \mu_{v_i\to f_j}(v_i)\, \mu_{f_j\to v_i}(v_i), \end{aligned} \tag{46}$$

where the last equality follows from (12). Comparing (45) and (46), we finally obtain (25).

APPENDIX C

In this appendix, we show that

$$p_{\mathbf{R},\mathbf{U}|\mathbf{B}}(\mathbf{r},\mathbf{u}|\mathbf{b}) \propto \gamma(\mathbf{b})\, p_{\mathbf{Y},\mathbf{U}|\mathbf{B}}(\mathbf{y}(\mathbf{b}),\mathbf{u}|\mathbf{b}), \tag{47}$$

where $\gamma(\mathbf{b})$ is defined in (33) and $\propto$ denotes equality up to a factor independent of $\mathbf{b}$.

We first note that[11]

$$\begin{aligned} p_{\mathbf{R},\mathbf{U}|\mathbf{B}}(\mathbf{r},\mathbf{u}|\mathbf{b}) &\triangleq \sum_{\mathbf{a},\mathbf{x}} p_{\mathbf{R},\mathbf{A},\mathbf{X},\mathbf{U}|\mathbf{B}}(\mathbf{r},\mathbf{a},\mathbf{x},\mathbf{u}|\mathbf{b}) \\ &= \sum_{\mathbf{a},\mathbf{x}} p_{\mathbf{R}|\mathbf{A},\mathbf{B}}(\mathbf{r}|\mathbf{a},\mathbf{b}) p_{\mathbf{A},\mathbf{X},\mathbf{U}}(\mathbf{a},\mathbf{x},\mathbf{u}), \end{aligned}$$

and

$$\begin{aligned} p_{\mathbf{Y},\mathbf{U}|\mathbf{B}}(\mathbf{y},\mathbf{u}|\mathbf{b}) &\triangleq \sum_{\mathbf{a},\mathbf{x}} p_{\mathbf{Y},\mathbf{A},\mathbf{X},\mathbf{U}|\mathbf{B}}(\mathbf{y},\mathbf{a},\mathbf{x},\mathbf{u}|\mathbf{b}) \\ &= \sum_{\mathbf{a},\mathbf{x}} p_{\mathbf{Y}|\mathbf{A},\mathbf{B}}(\mathbf{y}|\mathbf{a},\mathbf{b}) p_{\mathbf{A},\mathbf{X},\mathbf{U}}(\mathbf{a},\mathbf{x},\mathbf{u}), \end{aligned}$$

where

$$p_{\mathbf{Y}|\mathbf{A},\mathbf{B}}(\mathbf{y}|\mathbf{a},\mathbf{b}) = \left(\frac{1}{2N_0}\right)^K \exp\Big\{-\frac{\|\mathbf{y}-\mathbf{a}\|^2}{2N_0}\Big\},$$

is the distribution of the matched-filter outputs given $\mathbf{a}$ and $\mathbf{b}$.

Therefore, (47) is proved if we show that

$$p_{\mathbf{R}|\mathbf{A},\mathbf{B}}(\mathbf{r}|\mathbf{a},\mathbf{b}) \propto \gamma(\mathbf{b})\, p_{\mathbf{Y}|\mathbf{A},\mathbf{B}}(\mathbf{y}(\mathbf{b})|\mathbf{a},\mathbf{b}). \tag{48}$$

Now, due to the Gaussian nature of the noise affecting the observation samples $r(lT_s)$, we have

$$\begin{aligned} &p_{\mathbf{R}|\mathbf{A},\mathbf{B}}(\mathbf{r}|\mathbf{a},\mathbf{b}) \propto \\ &\exp\Big\{\frac{2\,\Re\{\sum_k a_k^* T_s \sum_l r(lT_s) c(lT_s - kT - k_\tau T)\, e^{-jk_\theta \Psi}\}}{2N_0}\Big\} \\ &\times \exp\Big\{-\frac{T_s\,\|\mathbf{r}\|^2}{2N_0}\Big\}\, \exp\Big\{-\frac{\|\mathbf{a}\|^2}{2N_0}\Big\}, \end{aligned} \tag{49}$$

where we used the fact that

$$T_s \sum_l c(lT_s - kT - k_\tau T) c(lT_s - k'T - k_\tau T) = \delta(k-k'), \tag{50}$$

by definition of root-raised cosine filters [33].

[11] We remind the reader that $\mathbf{x}$ stands for the sequence of coded bits.

Since $\exp\{-\frac{T_s\|\mathbf{r}\|^2}{2N_0}\}$ does not depend on $\mathbf{b}$, we can drop it. On the other hand, using the definition of the matched-filter output (31) and completing the square in (49), we have

$$p_{\mathbf{R}|\mathbf{A},\mathbf{B}}(\mathbf{r}|\mathbf{a},\mathbf{b}) \propto \exp\Big\{\frac{\|\mathbf{y}(\mathbf{b})\|^2}{2N_0}\Big\}\exp\Big\{-\frac{\|\mathbf{y}(\mathbf{b})-\mathbf{a}\|^2}{2N_0}\Big\}. \tag{51}$$

Comparing (51) to (48), we obtain the result.

REFERENCES

[1] J. Massey, "Optimum frame synchronization," *IEEE Trans. Commun.*, vol. 20, no. 2, pp. 115–119, Apr. 1972.

[2] G. Lui and H. Tan, "Frame synchronization for gaussian channels," *IEEE Trans. Commun.*, vol. 35, no. 8, pp. 818–829, Aug. 1987.

[3] E. Cacciamani and J. Wolejsza, C., "Phase-ambiguity resolution in a four-phase PSK communications system," *IEEE Trans. Commun. Technol.*, vol. 19, no. 6, pp. 1200–1210, Dec. 1971.

[4] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding," in *Proc. IEEE Int. Conf. on Commun.*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[5] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[6] W. Oh and K. Cheun, "Joint decoding and carrier phase recovery algorithm for turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 375–377, Sep. 2001.

[7] A. Nayak, J. Barry, and S. McLaughlin, "Joint timing recovery and turbo equalization for coded partial response channels," *IEEE Trans. Magn.*, vol. 38, no. 5, pp. 2295–2297, Sep. 2002.

[8] L. Zhang and A. Burr, "Application of turbo-principle to carrier phase recovery in turbo encoded bit-interleaved coded modulation system," in *Int. Symp. on Turbo Codes and Rel. Topics*, Brest, France, Sep. 2003, pp. 87–90.

[9] C. Herzet, H. Wymeersch, M. Moeneclaey, and L. Vandendorpe, "On maximum-likelihood timing synchronization," *IEEE Trans. Commun.*, vol. 55, no. 6, pp. 1116–1119, Jun. 2007.

[10] J. W. Walsh, J. C. R. Johnson, and P. A. Regalia, "Joint synchronization and decoding exploiting the turbo principle," in *Proc. Conf. on Inform. Sci. and Sys.*, Princeton, NJ, USA, Mar. 2004, pp. 17–19.

[11] J. Gunther, D. Keller, and T. Moon, "A generalized BCJR algorithm and its use in turbo synchronization," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, Logan, UT, USA, Mar. 2005, pp. 837–840.

[12] C. Herzet, V. Ramon, and L. Vandendorpe, "A theoretical framework for iterative synchronization based on the sum-product and the expectation-maximization algorithms," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1644–1658, May 2007.

[13] Q. Zhao and G. L. Stuber, "Turbo synchronization for serially concatenated CPM," *Proc. IEEE Int. Conf. on Commun.*, vol. 7, pp. 2976–2980, Jun. 2006.

[14] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, and L. Vandendorpe, "Code-aided turbo synchronization," *Proc. IEEE*, vol. 95, no. 6, pp. 1255–1271, Jun. 2007.

[15] U. Mengali, A. Sandri, and A. Spalvieri, "Phase ambiguity resolution in trellis-coded modulations," *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2087–2088, Dec. 1990.

[16] G. Lorden, R. McEliece, and L. Swanson, "Node synchronization for the viterbi decoder," *IEEE Trans. Commun.*, vol. 32, no. 5, pp. 524–531, May 1984.

[17] M. Moeneclaey and P. Sanders, "Syndrome-based viterbi decoder node synchronization and out-of-lock detection," in *Proc. IEEE Global Telecomm. Conf.*, vol. 1, Dallas, Texas, USA, Nov. 1990, pp. 604–608.

[18] B. Mielczarek and A. Svensson, "Timing error recovery in turbo-coded systems on AWGN channels," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1584–1592, Oct. 2002.

[19] T. M. Cassaro and C. N. Georghiades, "Frame synchronization for coded systems over AWGN channels," *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 484–489, Mar. 2004.

[20] U. Mengali, R. Pellizzoni, and A. Spalvieri, "Soft-decision-based node synchronization for viterbi decoders," *IEEE Trans. Commun.*, vol. 43, no. 9, pp. 2532–2539, Sep. 1995.

[21] H. Wymeersch and M. Moeneclaey, "Iterative code-aided ML phase estimation and phase ambiguity resolution," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 6, pp. 981–988, 2005.

[22] J. Dauwels, H. Wymeersch, and H.-A. Loeliger, "Phase estimation and phase ambiguity resolution by message passing," in *Int. Conf. on Telecommun. and Networking*, Fortaleza, Brazil, Aug. 2004, pp. 150–155.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, Jan. 1977.

[24] C. Herzet, H. Wymeersch, F. Simoens, M. Moeneclaey, and L. Vandendorpe, "MAP-based code-aided hypothesis testing," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 2856–2860, Aug. 2008.

[25] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

[26] M. Nissila and S. Pasupathy, "Adaptive iterative detectors for phase-uncertain channels via variational bounding," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 716–725, Mar. 2009.

[27] L. P. B. Christensen and J. Larsen, "On data and parameter estimation using the variational bayesian EM-algorithm for block-fading frequency-selective MIMO channels," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 4, Toulouse, France, May 2006, pp. 465–468.

[28] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1862–1874, May 2007.

[29] B. Hu, I. Land, L. Rasmussen, R. Piton, and B. H. Fleury, "A divergence minimization approach to joint multiuser decoding for coded CDMA," *IEEE J. Select. Areas Commun.*, vol. 26, no. 3, pp. 432–445, Apr. 2008.

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

[31] P. O. Vontobel, "A factor-graph approach to universal decoding," in *Allerton Conf. on Commun., Contr., and Comput.*, Monticello, IL, USA, Sep. 2006.

[32] C. Herzet, "On the convergence of the iterative "pseudo likelihood" maximization algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, USA, Mar. 2008, pp. 3701–3704.

[33] J. Proakis, *Digital Communications*, 3rd ed. McGraw-Hill, 1995.

[34] H. Meyr, M. Moenenclaey, and S. Fetchel, *Digital Communication Receivers : Synchronization, Channel Estimation and Signal Processing*, ser. Telecommunications and Signal Processing. USA: Wiley, 1998.

[35] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, ser. Applications of Communications Theory. NY, USA: Springer, 1997.

[36] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Telecomm. Conf.*, vol. 1, Sidney, Australia, Nov. 1998, pp. 579–584.

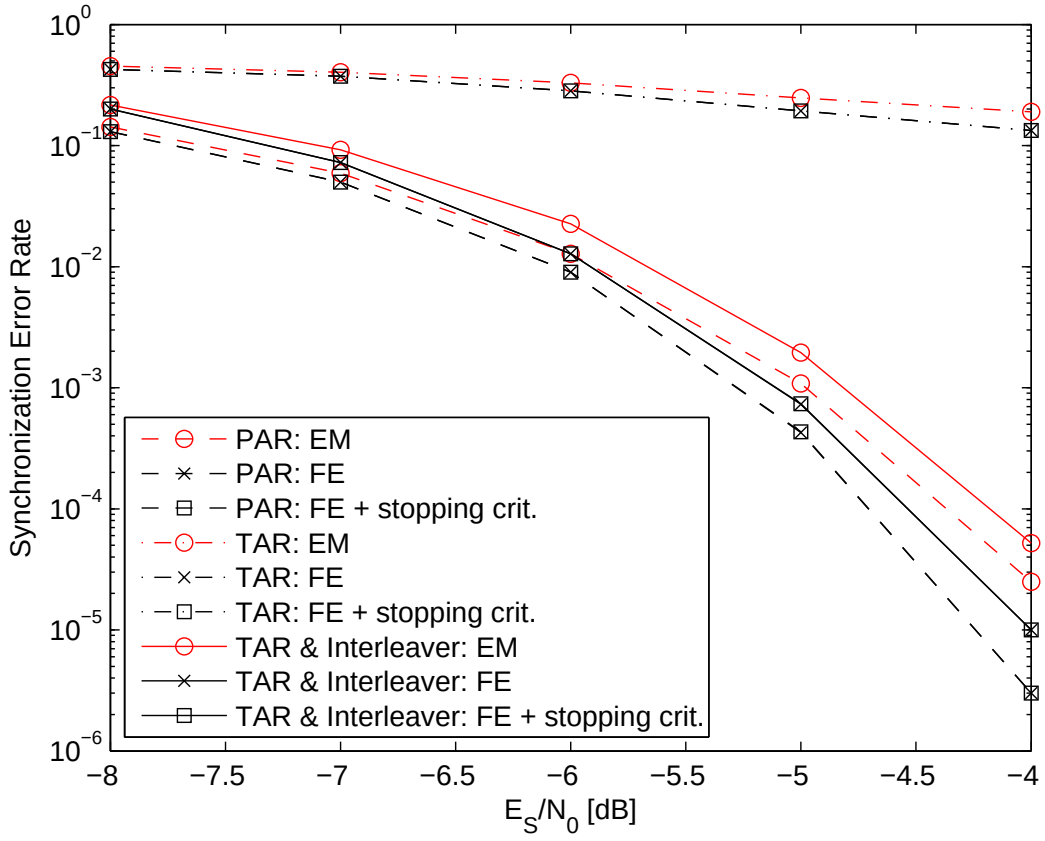

Figure 2. Synchronization error rate of timing ambiguity resolution (TAR) and phase ambiguity resolution (PAR) for a convolutionally-coded transmission.

[37] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[38] S. Aji and R. McEliece, "The generalized distributive law," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 325–343, Mar. 2000.

[39] A. L. Yuille, "CCCP algorithms to minimize the Bethe and Kikuchi free energies: Convergent alternatives to belief propagation," *Neural Computation*, vol. 14, pp. 1691–1722, Jul. 2002.

[40] M. Welling and Y. W. Teh, "Belief optimization for binary networks: A stable alternative to loopy belief propagation," in *Conf. on Uncertainty in Artificial Intell.*, San Francisco, CA, USA, Jul. 2001, pp. 554–561.

[41] H. J. Kappen and W. Wiegerinck, "Novel iteration schemes for the cluster variation method," in *Advances in Neural Inform. Processing Syst.*, vol. 14, Vancouver, BC, Canada, Dec. 2001, pp. 415–422.

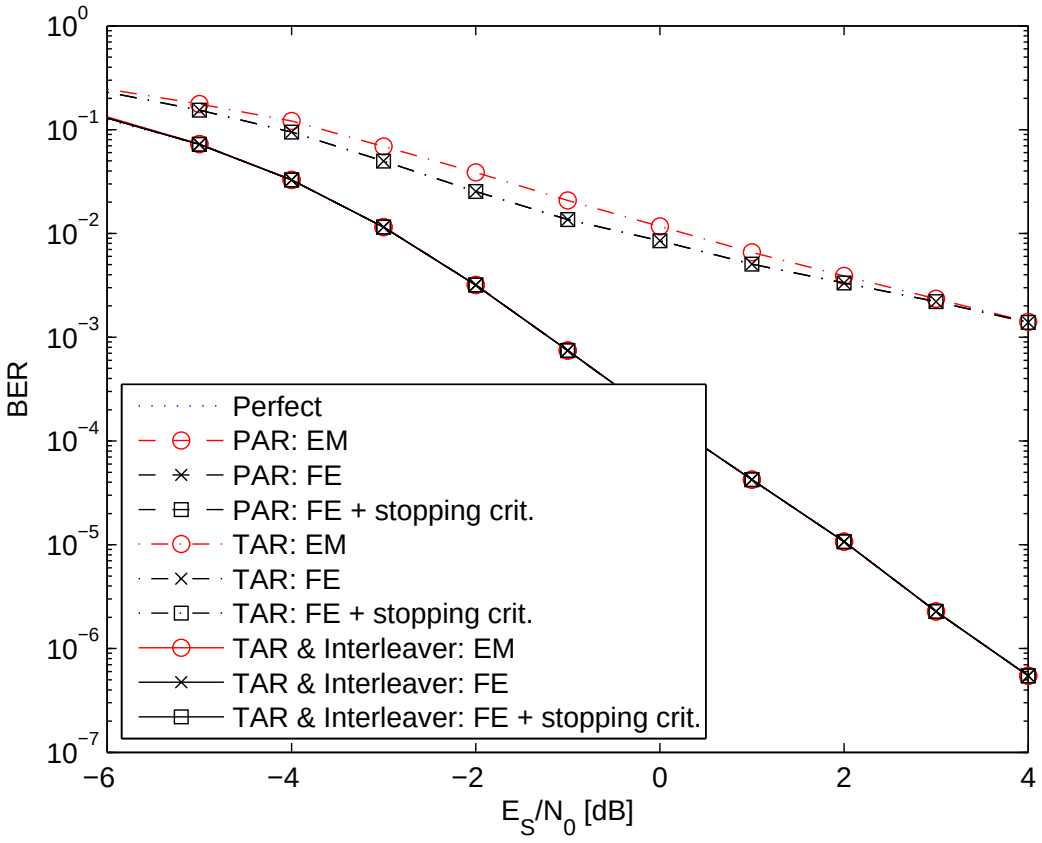

Figure 3. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) algorithms for a convolutionally-coded transmission.

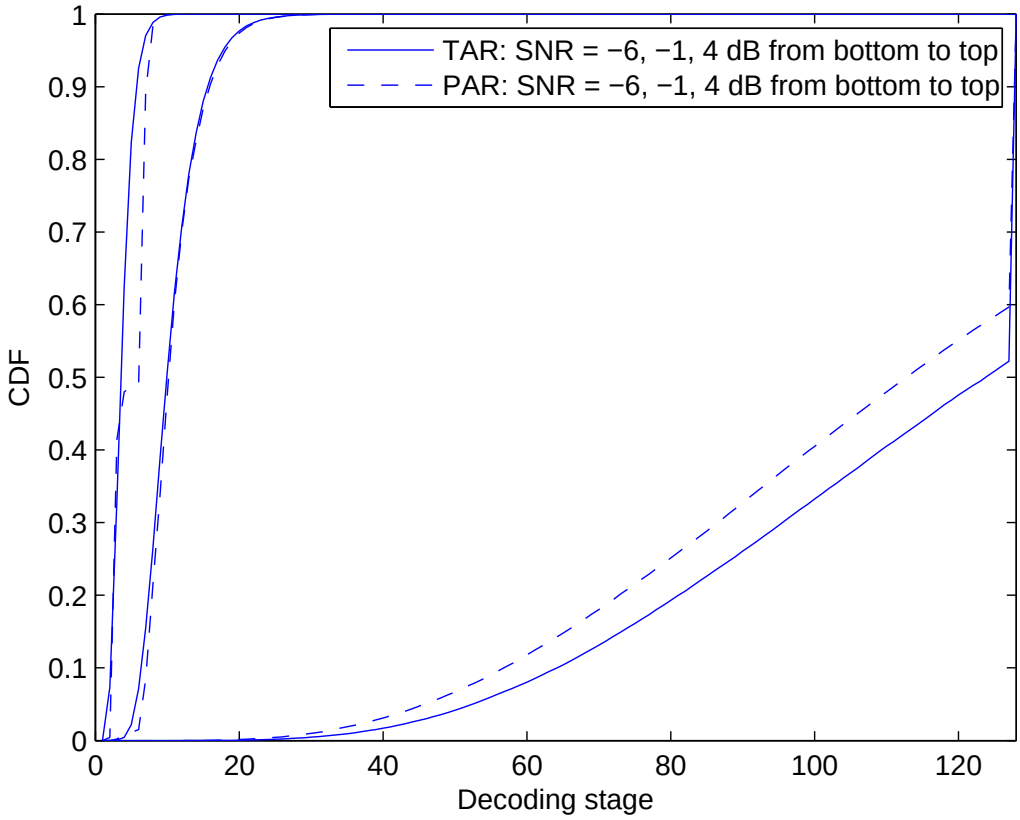

Figure 4. CDF of the decoding stage at which the free energy minimization (FE) algorithm with stopping criterion makes a decision for phase ambiguity resolution (PAR) and timing ambiguity resolution (TAR) with interleaver in the case of convolutionally-coded transmission.

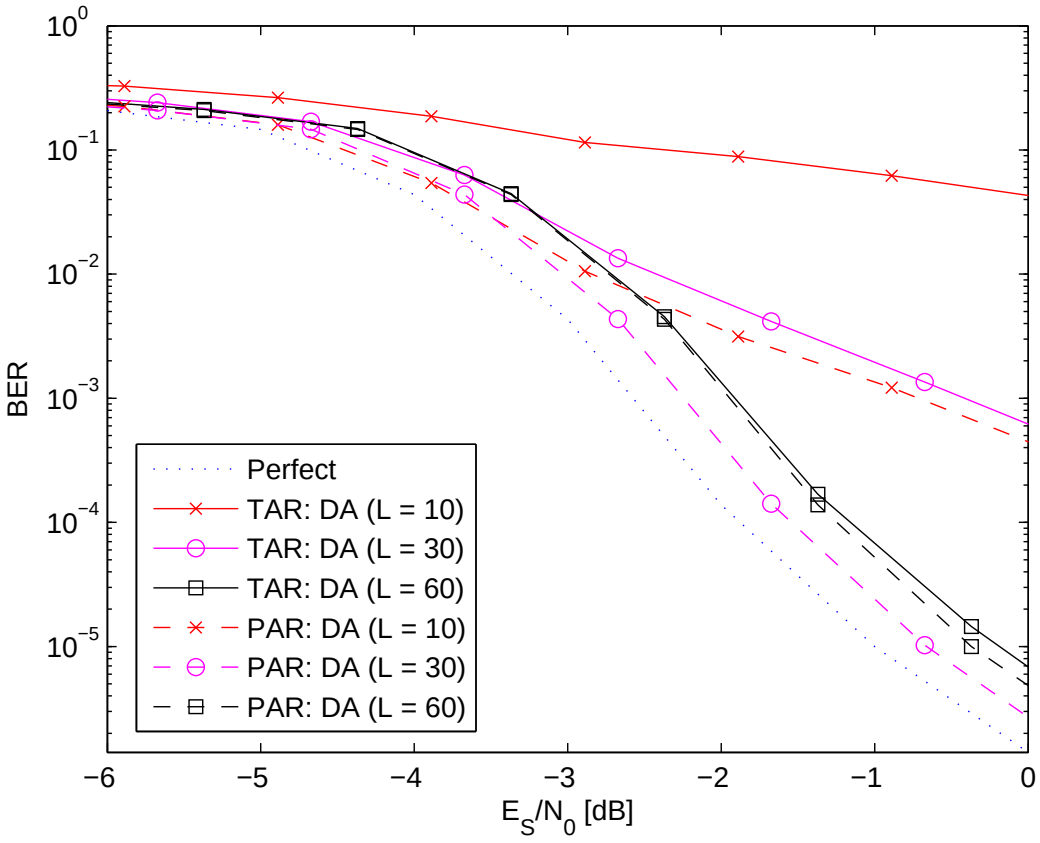

Figure 5. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) with data-aided (DA) algorithm at a pilot length of 10, 30 and 60.

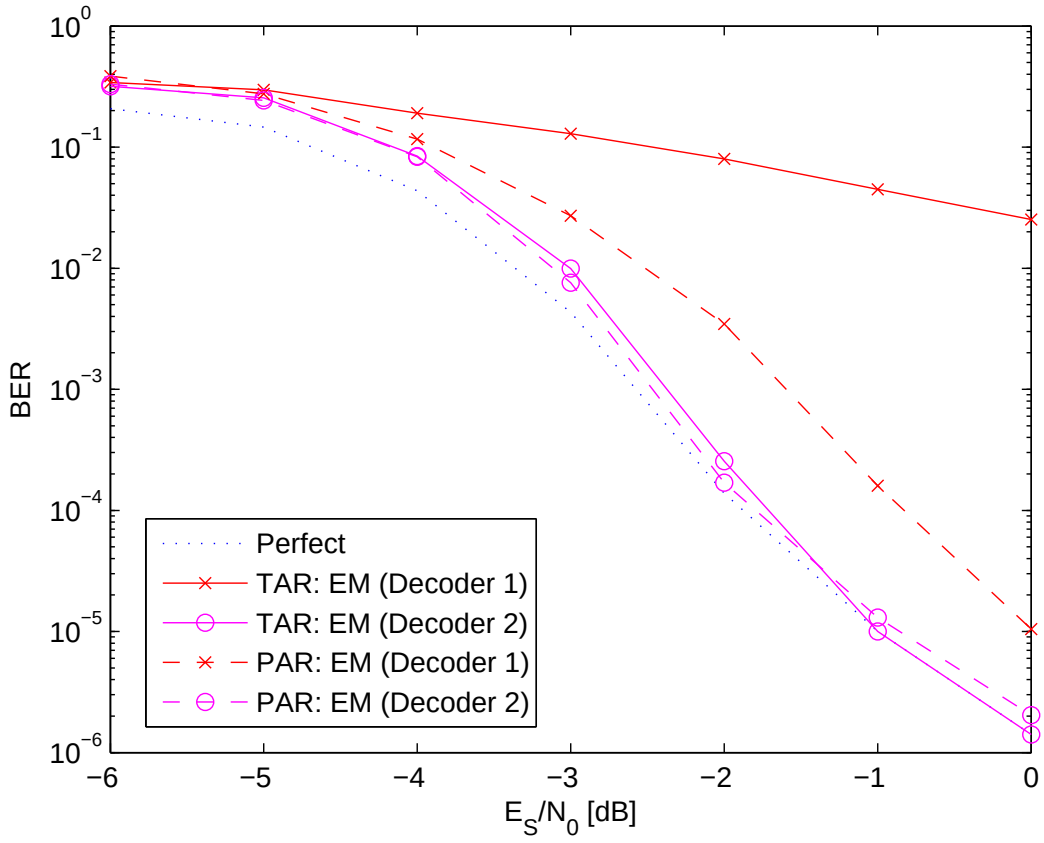

Figure 6. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) with expectation-maximization (EM) algorithm making decision at decoder 1 or at decoder 2. The length of the coded sequence is 128.

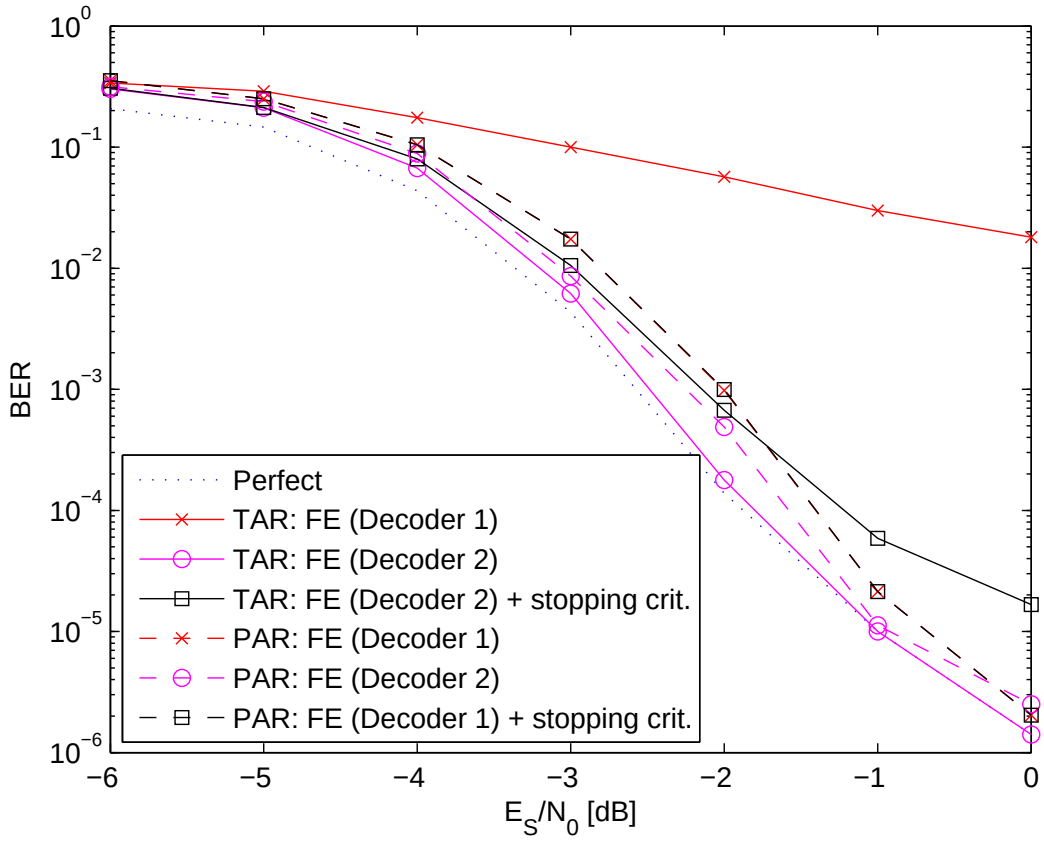

Figure 7. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) with the free energy minimization (FE) algorithm making decision at decoder 1 or at decoder 2. The length of the coded sequence is 128.

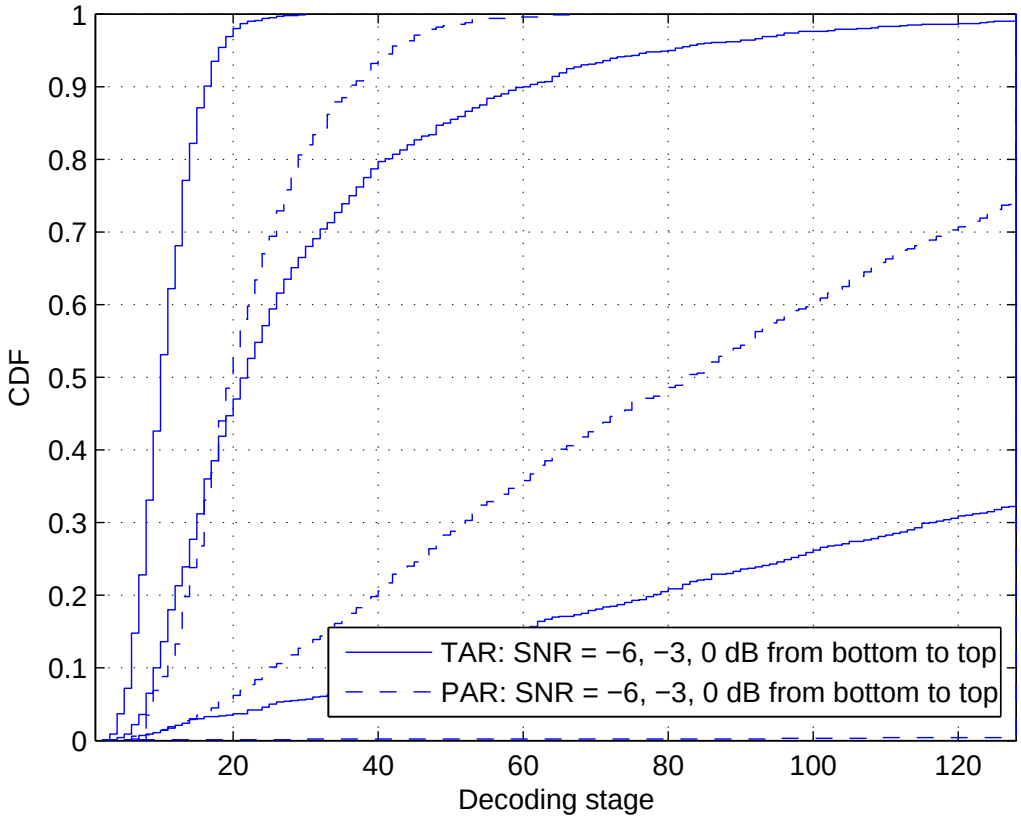

Figure 8. CDF of the decoding stage at which the free energy minimization (FE) algorithm with stopping criterion makes a decision at decoder 1 for phase ambiguity resolution (PAR) and at decoder 2 for timing ambiguity resolution (TAR). The length of the coded sequence is 128.

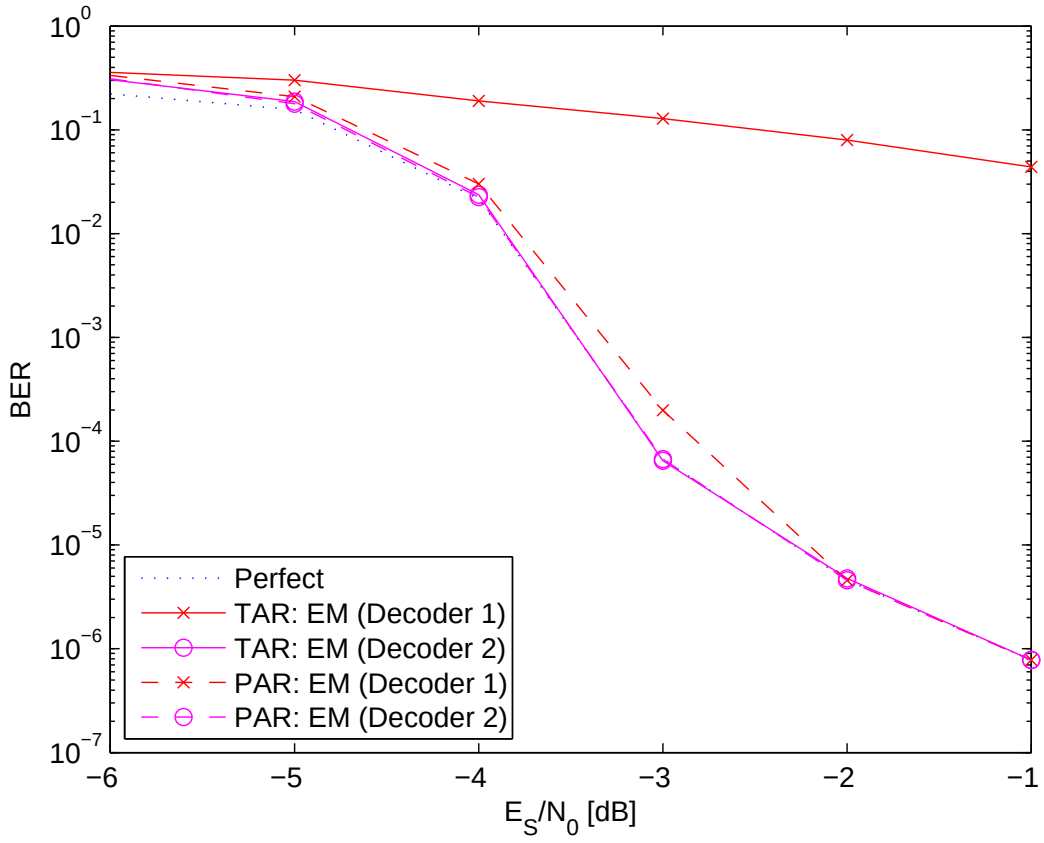

Figure 9. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) with expectation-maximization (EM) algorithm making decision at decoder 1 or at decoder 2. The length of the coded sequence is 512.

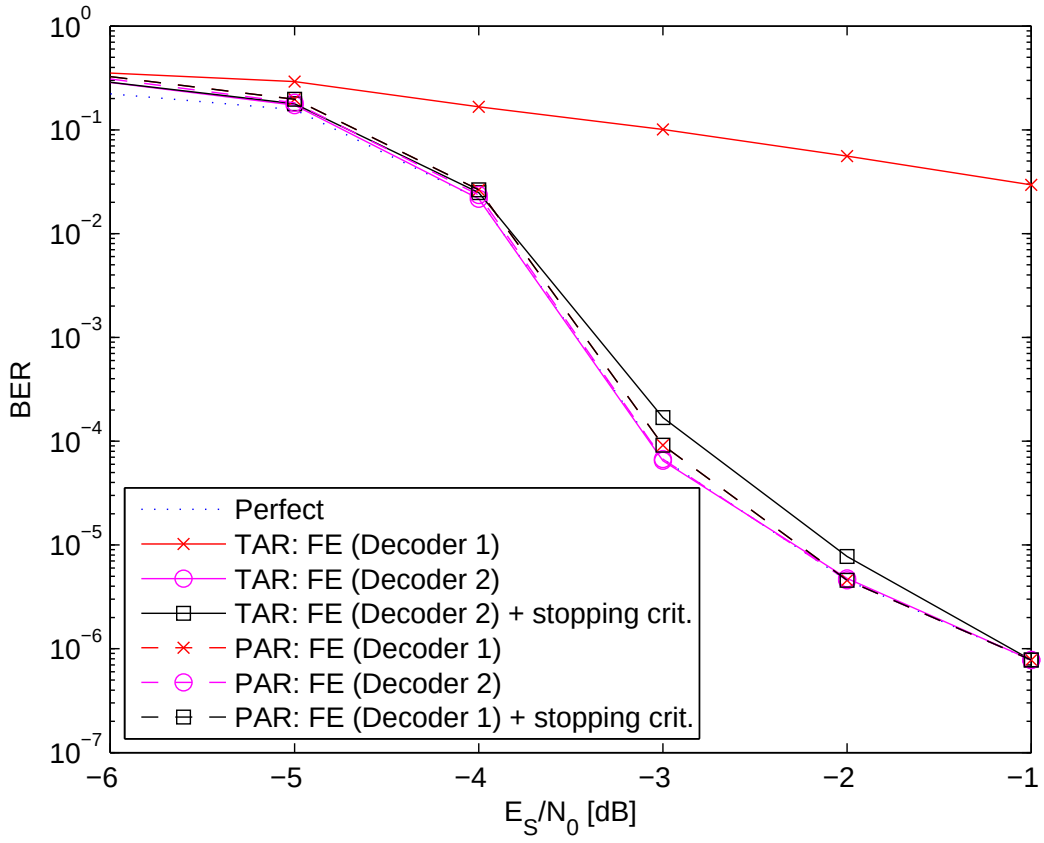

Figure 10. BER comparison between perfect synchronization and timing ambiguity resolution (TAR) or phase ambiguity resolution (PAR) with the free energy minimization (FE) algorithm making decision at decoder 1 or at decoder 2. The length of the coded sequence is 512.