Iterative Receiver Design for ISI Channels Using Combined Belief- and Expectation-Propagation

Sun, Peng; Zhang, Chuanzong; Wang, Zhongyong; Manchón, Carles Navarro; Fleury, Bernard Henri

Published in:
I E E E Signal Processing Letters

DOI (link to publication from Publisher):
10.1109/LSP.2015.2404822

Publication date:
2015

Document Version
Peer reviewed version

Iterative Receiver Design for ISI Channels Using Combined Belief- and Expectation-Propagation

Peng Sun, Chuanzong Zhang, Zhongyong Wang, Carles N. Manchón, and Bernard H. Fleury, *Senior member, IEEE*

***Abstract*—In this letter, a message-passing algorithm that combines belief propagation and expectation propagation is applied to design an iterative receiver for intersymbol interference channels. We detail the derivation of the messages passed along the nodes of a vector-form factor graph representing the underlying probabilistic model. We also present a simple but efficient method to cope with the "negative variance" problem of expectation propagation. Simulation results show that the proposed algorithm outperforms, in terms of bit-error-rate and convergence rate, a LMMSE turbo-equalizer based on Gaussian message passing with the same order of computational complexity.**

***Index Terms*—belief propagation, expectation propagation, turbo equalization.**

I. Introduction

SINCE optimal detection of data transmitted across an intersymbol interference (ISI) channel, like the multipath wireless channel, is typically impractical, suboptimal receiver structures that approach the performance of the optimal detector have been proposed, with turbo equalization [1] being the most emblematic instance. In turbo equalization, the turbo principle –originally used for decoding concatenated codes [2]– is applied by regarding the ISI channel as an encoder acting on the transmitted symbols.

The above "turbo"-processing algorithms are instances of belief propagation (BP) applied on a factor graph representing the underlying probabilistic model [3]. Additional equalizer structures, which implement other variants of BP, have been proposed, e.g. [4]. However, BP-based equalizers suffer from an inherent drawback: their complexity grows exponentially with the channel length or the number of non-zero coefficients (depending on the selected factor graph representation) and the modulation order.

Manuscript received December 04, 2014; revised January 22, 2015; accepted January 22, 2015. Date of publication nulldate; date of current version nulldate. This work was conducted in cooperation with the 4GMCT cooperative research project funded by Intel Mobile Communications Denmark ApS, Agilent Technologies, Aalborg University, and the Danish National Advanced Technology Foundation. This work was supported by the National Natural Science Foundation of China under NSFC Grants 61172086, 61201251, and 61401401, and also by the National 863 Program of China under Grant 0214AA01A705. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rodrigo C. de Lamare.

P. Sun and Z. Wang are with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China (e-mail: iepengsun@gmail.com; iezywang@zzu.edu.cn).

C. Zhang is with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China, and the Department of Electronic Systems, Aalborg University, Aalborg 9220, Denmark (e-mail: ieczzhang@gmail.com).

C. N. Manchón and B. H. Fleury are with the Department of Electronic Systems, Aalborg University, Aalborg 9220, Denmark (e-mail: cnm@es.aau.dk; fleury@es.aau.dk).

Different approaches have been proposed to circumvent the aforementioned complexity issue. Basically, they introduce approximations that make the messages passed in the subgraph representing the ISI channel Gaussian. In [5] this is achieved by assuming the interference plus noise component with respect to each modulation symbol to be Gaussian and exploiting a relationship between the extrinsic values of the symbols when the channel is driven by these symbols and the LMMSE symbol estimates when the channel is driven by Gaussian inputs. This approach turns out to be equivalent to that in [6] when applied to turbo-equalization [5]. In [7] a combined use of Gaussian expectation propagation (EP) [8], [9] and BP is proposed. The use of EP, however, leads to an unstable algorithm due to the fact that computed Gaussian EP messages may have a negative variance. In [7] the authors propose to circumvent this problem by replacing each EP message with a geometric mixture of said message and a standard Gaussian message, parameterized with a damping/mixing factor. However, "good" sequences of values of the damping factor versus the iteration index of the algorithm need to be tuned in advance via simulations, which severely limits the practicability of the proposed approach.

In this paper we formulate an approximate inference method combining BP and EP and apply it to a vector-form factor graph representation of the probabilistic model for ISI channels to design a receiver algorithm performing joint equalization of ISI channels and data detection. The obtained design is similar to that presented in [7]. We propose a simple solution to avoid the instability problem of EP that leads to a fast converging algorithm. We present a detailed derivation of the turbo-equalizer and a numerical evaluation that compares its performance with that of the receiver proposed in [5]. The simulation results show that for the same complexity our design performs better and converges faster than that in [5], while avoiding the practical issues inherent to that in [7].

Notation- Boldface lowercase and uppercase letters denote vectors and matrices, respectively. The identity matrix of size M is represented by $\boldsymbol{I}_M$. Superscript $(\cdot)^{\mathrm{T}}$ indicates transposition of a vector or matrix. The probability density function (pdf) of a multivariate Gaussian distribution with mean vector $\boldsymbol{m}$ and covariance matrix $\boldsymbol{V}$ is represented by $\mathcal{N}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{V})$. The relation $f(x)=cg(x)$ for some positive constant c is written as $f(x)\propto g(x)$.

II. System Model

The information bit vector $\boldsymbol{b}=[b_1,\ldots,b_K]^{\mathrm{T}}$ is encoded and interleaved, yielding the codeword vector $\boldsymbol{c}=[c_1,\ldots,c_N]^{\mathrm{T}}$.

The coded bits are then mapped onto a binary phase shift keying (BPSK) constellation, resulting in the vector of modulated symbols $\boldsymbol{x} = [x_1, \ldots, x_N]^{\mathrm{T}}$, which are then transmitted over a frequency-selective channel corrupted with AWGN. The (baseband discrete-time) signal observed at the receiver is described by the vector $\boldsymbol{r} = [r_1, \ldots, r_{N+L-1}]^{\mathrm{T}}$ with entries

$$r_i = \sum_{l=0}^{L-1} h_l x_{i-l} + n_i = \boldsymbol{h}^{\mathrm{T}} \boldsymbol{s}_i + n_i. \quad (1)$$

Here, $\boldsymbol{s}_i = [x_{i-L+1}, \ldots, x_i]^{\mathrm{T}}$ with $x_i = 0$ for $i < 1$ and $i > N$, $\boldsymbol{h} = [h_{L-1}, \ldots, h_0]^{\mathrm{T}}$ denotes the vector of channel weights, and n_i is the ith sample of a white Gaussian noise vector with component variance σ^2.

A. Probabilistic Model and Factor Graph

The posterior probability mass function (pmf) of vectors $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{x}$ and $\boldsymbol{s}$ given the received signal vector $\boldsymbol{r}$ reads

$$\begin{aligned} p(\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{x}, \boldsymbol{s} | \boldsymbol{r}) \propto & \prod_{k=1}^{K} f_{b_k}(b_k) \times f_c(\boldsymbol{c}, \boldsymbol{b}) \\ & \times \prod_{i=1}^{N} f_{r_i}(r_i, \boldsymbol{s}_i) f_{G_i}(\boldsymbol{s}_i, \boldsymbol{s}_{i-1}, x_i) f_{M_i}(x_i, c_i) \\ & \times \prod_{i=N+1}^{N+L-1} f_{r_i}(r_i, \boldsymbol{s}_i) f_{G_i}(\boldsymbol{s}_i, \boldsymbol{s}_{i-1}, 0). \quad (2) \end{aligned}$$

In this expression $f_{b_k}(b_k)$ is the uniform prior pmf of the kth information bit, $f_c(\boldsymbol{c}, \boldsymbol{b})$ stands for the coding and interleaving constraints, $f_{r_i}(r_i, \boldsymbol{s}_i) \triangleq p(r_i|\boldsymbol{s}_i) \propto \mathcal{N}(r_i; \boldsymbol{h}^{\mathrm{T}}\boldsymbol{s}_i, \sigma^2)$ denotes the likelihood term for $\boldsymbol{s}_i$, and $f_{M_i}(x_i, c_i)$ represents the modulation mapping. Finally, $f_{G_i}(\boldsymbol{s}_i, \boldsymbol{s}_{i-1}, x_i)$ expresses the deterministic relationship between $\boldsymbol{s}_i$, $\boldsymbol{s}_{i-1}$ and x_i, given by

$$\boldsymbol{s}_i = \boldsymbol{G}\boldsymbol{s}_{i-1} + \boldsymbol{e}x_i \quad (3)$$

with the $L \times L$ matrix $\boldsymbol{G} = [\boldsymbol{0} \quad \boldsymbol{I}_{L-1}; \quad 0 \quad \boldsymbol{0}^{\mathrm{T}}]$ and the L vector $\boldsymbol{e} = [\boldsymbol{0}^{\mathrm{T}} \quad 1]^{\mathrm{T}}$, where $\boldsymbol{0}$ is a zero column vector with length $L-1$. Note that $\boldsymbol{G}$ factorizes as $\boldsymbol{G} = \boldsymbol{G}''\boldsymbol{G}'$ with $\boldsymbol{G}'' = [\ \boldsymbol{I}_{L-1} \quad \boldsymbol{0}\ ]^{\mathrm{T}}$ and $\boldsymbol{G}' = [\ \boldsymbol{0} \quad \boldsymbol{I}_{L-1}\ ]$ [5].

The vector-form factor graph representation [5] of the posterior pmf in (2) is depicted in Fig. 1. It will be used for the derivation of the BP-EP-based receiver described in Section IV. Note that in this representation the subgraph representing the ISI channel (left part) in Fig. 1 has a tree structure[1].

III. Combined BP-EP message-passing rule

We consider a factor graph of a generic probabilistic model made of a set of factor nodes $\mathcal{F}$, and a set of variable nodes $\mathcal{Z}$. The variable nodes are grouped into two disjoint subsets $\mathcal{Z}^{\mathrm{BP}}$ and $\mathcal{Z}^{\mathrm{EP}}$, i.e. $\mathcal{Z}^{\mathrm{BP}} \cup \mathcal{Z}^{\mathrm{EP}} = \mathcal{Z}$ and $\mathcal{Z}^{\mathrm{BP}} \cap \mathcal{Z}^{\mathrm{EP}} = \emptyset$. Let $m_{f\to z}(z)$ denote the messages from a factor node $f \in \mathcal{F}$ to a variable node $z \in \mathcal{Z}$, and $n_{z\to f}(z)$ be the message from variable node z to factor node f. With these definitions, the message update rules read

$$m_{f\to z}(z) = \sum_{\sim\{z\}} f(\boldsymbol{z}) \prod_{z' \in \mathcal{N}(f)\setminus\{z\}} n_{z'\to f}(z'), z \in \mathcal{Z}^{\mathrm{BP}} \quad (4)$$

$$m_{f\to z}(z) = \frac{\mathrm{Proj}_{\mathcal{E}_z}\left[m^{\mathrm{BP}}_{f\to z}(z)\, n_{z\to f}(z)\right]}{n_{z\to f}(z)}, \quad z \in \mathcal{Z}^{\mathrm{EP}} \quad (5)$$

$$n_{z\to f}(z) = \prod_{f' \in \mathcal{N}(z)\setminus\{f\}} m_{f'\to z}(z), \quad z \in \mathcal{Z}. \quad (6)$$

Here, $\sum_{\sim\{z\}}$ is the sum over all variables of $f = f(\boldsymbol{z})$ excluding z, $\mathcal{N}(z)$ and $\mathcal{N}(f)$ denote respectively the set of factor nodes connected to variable node z and the set of variable nodes connected to the factor node f. The superscript BP of the message in the right-hand expression in (5) indicates that this message from factor f to $z \in \mathcal{Z}^{\mathrm{EP}}$ is computed using the BP rule, i.e. (4). Moreover in this expression $\mathrm{Proj}_{\mathcal{E}_z}[\cdot]$ is the projection of the pdf given as an argument on a specified exponential family $\mathcal{E}_z$ [2]. Note that computing the messages from any factor node to any variable node requires the computation of a BP message. Moreover, messages passed from and to a variable node $z \in \mathcal{Z}^{\mathrm{BP}}$ ($z \in \mathcal{Z}^{\mathrm{EP}}$) are computed using the BP (EP) rule.

IV. Iterative Receiver Design

In this section we derive a receiver that performs joint equalization and decoding for ISI channels by passing messages along the edges of the factor graph depicted in Fig. 1. The complexity of standard BP applied on this factor graph grows exponentially with L, the dimension of the state vectors $\boldsymbol{s}_i$, $\forall i$. Such intractable complexity can be reduced by approximating messages passed along the edges of the channel part of the graph (to the left of and including the variable nodes x_i, $\forall i$) with Gaussian messages. This can be done by approximating the messages from x_i to f_{G_i}, $\forall i$, with Gaussian messages. The EP framework provides an elegant and efficient tool to do so. Similarly to [7] we split the variable nodes in the graph as follows: $\mathcal{Z}^{\mathrm{EP}} = \{x_i; \forall i\}$ and $\mathcal{Z}^{\mathrm{BP}} = \mathcal{Z} \setminus \mathcal{Z}^{\mathrm{EP}}$. Moreover we set $\mathcal{E}_{x_i} = \mathcal{G}$, $\forall i$, where $\mathcal{G}$ is the Gaussian family.

A. Calculation of messages

1) Equalization – input messages: Assuming that, for the ith symbol, the message from f_{M_i} to x_i can

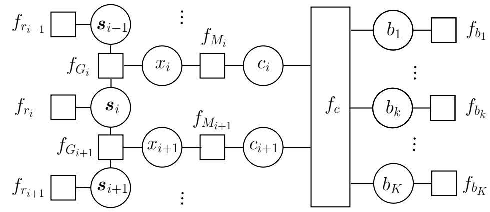

Fig. 1. Vector-form factor graph representation of the probabilistic model (2).

[1]Cycles appearing in "channel" subgraph of the scalar-form factor graph representation of the probabilistic model (2) are absorbed in the vector-form representation [5].

[2]For a pdf $b(z)$ $\mathrm{Proj}_{\mathcal{E}}[b(z)] = \arg\min_{b\prime(z)\in\mathcal{E}} D\left(b(z)||b\prime(z)\right)$, with $D(\cdot||\cdot)$ denoting the Kullback-Leibler divergence, and $\mathcal{E}$ being a specific exponential family.
As indicated by the indexing, $\mathcal{E}_z$ might depend on the variable node.

be expressed as $m^{\mathrm{BP}}_{f_{M_i}\to x_i}(x_i) = \beta_{i,1}\delta(x_i+1) + \beta_{i,2}\delta(x_i-1)$ and the message from x_i to f_{M_i} has the form $n_{x_i\to f_{M_i}}(x_i) \propto \mathcal{N}(x_i;\vec{m}_{x_i},\vec{v}_{x_i})$, the belief $b(x_i) \propto m^{\mathrm{BP}}_{f_{M_i}\to x_i}(x_i)\, n_{x_i\to f_{M_i}}(x_i)$ of x_i has mean and variance

$$m^p_{x_i} = \frac{\beta_{i,2}\exp\{2\vec{m}_{x_i}/\vec{v}_{x_i}\} - \beta_{i,1}}{\beta_{i,2}\exp\{2\vec{m}_{x_i}/\vec{v}_{x_i}\} + \beta_{i,1}}, \quad (7)$$

$$v^p_{x_i} = 1 - (m^p_{x_i})^2. \quad (8)$$

The message $m_{f_{M_i}\to x_i}(x_i)$ is computed from (5) to be

$$\begin{aligned} m_{f_{M_i}\to x_i}(x_i) &= \frac{\mathrm{Proj}_{\mathcal{G}}[m^{\mathrm{BP}}_{f_{M_i}\to x_i}(x_i)\, n_{x_i\to f_{M_i}}(x_i)]}{n_{x_i\to f_{M_i}}(x_i)} \\ &\propto \mathcal{N}(x_i;\overleftarrow{m}_{x_i},\overleftarrow{v}_{x_i}) \end{aligned} \quad (9)$$

where

$$\overleftarrow{v}_{x_i} = [(v^p_{x_i})^{-1} - \vec{v}^{-1}_{x_i}]^{-1} \quad (10)$$

$$\overleftarrow{m}_{x_i} = \overleftarrow{v}_{x_i}[(v^p_{x_i})^{-1} m^p_{x_i} - \vec{v}^{-1}_{x_i}\vec{m}_{x_i}]. \quad (11)$$

We further have $n_{x_i\to f_{G_i}}(x_i) = m_{f_{M_i}\to x_i}(x_i)$ by (6).

When running the BP-EP algorithm, it can be observed that the variance parameter $\overleftarrow{v}_{x_i}$ in (10) and (11) sometimes takes negative values, which results in a bad performance, see also [7]. To avoid this problem, the variance $\overleftarrow{v}_{x_i}$ is replaced by its absolute value $|\overleftarrow{v}_{x_i}|$ in both (10) and (11). We will see in the numerical evaluations that this simple "trick" is very efficient and provides a viable alternative to the damping method proposed in [7], see Section I.

2) Equalization – downward messages: Assuming that the message $n_{\boldsymbol{s}_{i-1}\to f_{G_i}}(\boldsymbol{s}_{i-1}) \propto \mathcal{N}(\boldsymbol{s}_{i-1};\boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_{i-1}},\boldsymbol{V}^{\downarrow}_{\boldsymbol{s}_{i-1}})$ is known, the message $m_{f_{G_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i)$ is obtained via (4) to be

$$m_{f_{G_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{s}_i - \boldsymbol{m}^{\Downarrow}_{\boldsymbol{s}_i})^{\mathrm{T}}\boldsymbol{V}^{\Downarrow -1}_{\boldsymbol{s}_i}(\boldsymbol{s}_i - \boldsymbol{m}^{\Downarrow}_{\boldsymbol{s}_i})\right\} \quad (12)$$

with

$$\boldsymbol{m}^{\Downarrow}_{\boldsymbol{s}_i} = \boldsymbol{G}\boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_{i-1}} + \boldsymbol{e}\overleftarrow{m}_{x_i} \quad (13)$$

$$\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i} = \boldsymbol{G}\boldsymbol{V}^{\downarrow}_{\boldsymbol{s}_{i-1}}\boldsymbol{G}^{\mathrm{T}} + \boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}\overleftarrow{v}_{x_i}. \quad (14)$$

The message $n_{\boldsymbol{s}_i\to f_{G_{i+1}}}(\boldsymbol{s}_i)$ is calculated from (6) to be

$$\begin{aligned} n_{\boldsymbol{s}_i\to f_{G_{i+1}}}(\boldsymbol{s}_i) &= m_{f_{G_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i) m_{f_{r_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i) \quad (15) \\ &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{s}_i - \boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_i})^{\mathrm{T}}\boldsymbol{V}^{\downarrow -1}_{\boldsymbol{s}_i}(\boldsymbol{s}_i - \boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_i})\right\} \end{aligned}$$

where

$$\boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_i} = \boldsymbol{m}^{\Downarrow}_{\boldsymbol{s}_i} + \tfrac{1}{\sigma^2 + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i}\boldsymbol{h}}(r_i - \boldsymbol{h}^{\mathrm{T}}\boldsymbol{m}_i{}^{\Downarrow})\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i}\boldsymbol{h} \quad (16)$$

$$\boldsymbol{V}^{\downarrow}_{\boldsymbol{s}_i} = \boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i} - \tfrac{1}{\sigma^2 + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i}\boldsymbol{h}}\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i}\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}\boldsymbol{V}^{\Downarrow}_{\boldsymbol{s}_i} \quad (17)$$

and $m_{f_{r_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i) = f_{r_i}(r_i,\boldsymbol{s}_i)$.

3) Equalization – upward messages: With the message from variable node $\boldsymbol{s}_{i+1}$ to factor node $f_{G_{i+1}}$ being of the form $n_{\boldsymbol{s}_{i+1}\to f_{G_{i+1}}}(\boldsymbol{s}_{i+1}) \propto \mathcal{N}(\boldsymbol{s}_{i+1};\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_{i+1}},\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_{i+1}})$, the message $m_{f_{G_{i+1}}\to\boldsymbol{s}_i}(\boldsymbol{s}_i)$ from $f_{G_{i+1}}$ to $\boldsymbol{s}_i$ is obtained as

$$\begin{aligned} &m_{f_{G_{i+1}}\to\boldsymbol{s}_i}(\boldsymbol{s}_i) \propto \\ &\quad \exp\left\{-\tfrac{1}{2}(\boldsymbol{G}\boldsymbol{s}_i - \boldsymbol{m}^{\Uparrow}_{\boldsymbol{s}_i})^{\mathrm{T}}\boldsymbol{V}^{\Uparrow -1}_{\boldsymbol{s}_i}(\boldsymbol{G}\boldsymbol{s}_i - \boldsymbol{m}^{\Uparrow}_{\boldsymbol{s}_i})\right\} \end{aligned} \quad (18)$$

with

$$\begin{aligned} \boldsymbol{V}^{\Uparrow -1}_{\boldsymbol{s}_i}\boldsymbol{m}^{\Uparrow}_{\boldsymbol{s}_i} &= -\frac{\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{e}(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_{i+1}} + \overleftarrow{v}^{-1}_{x_{i+1}}\overleftarrow{m}_{x_{i+1}})}{\overleftarrow{v}^{-1}_{x_{i+1}} + \boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{e}} \\ &\quad + \boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_{i+1}} \end{aligned} \quad (19)$$

$$\boldsymbol{V}^{\Uparrow -1}_{\boldsymbol{s}_i} = \boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}} - \frac{\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}}{\overleftarrow{v}^{-1}_{x_{i+1}} + \boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_{i+1}}\boldsymbol{e}}. \quad (20)$$

As a consequence, the message $n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i)$ reads

$$n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{s}_i - \boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i})^{\mathrm{T}}\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_i}(\boldsymbol{s}_i - \boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i})\right\} \quad (21)$$

where

$$\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_i}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i} = \boldsymbol{G}^{\mathrm{T}}\boldsymbol{V}^{\Uparrow -1}_{\boldsymbol{s}_i}\boldsymbol{m}^{\Uparrow}_{\boldsymbol{s}_i} + \frac{\boldsymbol{h}r_i}{\sigma^2} \quad (22)$$

$$\boldsymbol{V}^{\uparrow -1}_{\boldsymbol{s}_i} = \boldsymbol{G}^{\mathrm{T}}\boldsymbol{V}^{\Uparrow -1}_{\boldsymbol{s}_i}\boldsymbol{G} + \frac{\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}}{\sigma^2}. \quad (23)$$

4) Equalization – output messages: The message $m^{\mathrm{BP}}_{f_{G_i}\to x_i}(x_i)$ reads

$$m^{\mathrm{BP}}_{f_{G_i}\to x_i}(x_i) \propto \exp\left\{-\frac{(x_i - \vec{m}_{x_i})^2}{2\vec{v}_{x_i}}\right\} \quad (24)$$

with

$$\begin{aligned} \vec{m}_{x_i} &= \boldsymbol{e}^{\mathrm{T}}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i} + \boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{G}''\left[\boldsymbol{G}'\boldsymbol{V}^{\downarrow}_{\boldsymbol{s}_{i-1}}\boldsymbol{G}'^{\mathrm{T}} + \boldsymbol{G}''^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{G}''\right]^{-1} \\ &\quad \times \left[\boldsymbol{G}'\boldsymbol{m}^{\downarrow}_{\boldsymbol{s}_{i-1}} - \boldsymbol{G}''^{\mathrm{T}}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i}\right] \end{aligned} \quad (25)$$

$$\begin{aligned} \vec{v}_{x_i} &= \boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{e} - \boldsymbol{e}^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{G}''\left[\boldsymbol{G}'\boldsymbol{V}^{\downarrow}_{\boldsymbol{s}_{i-1}}\boldsymbol{G}'^{\mathrm{T}} + \boldsymbol{G}''^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{G}''\right]^{-1} \\ &\quad \times \boldsymbol{G}''^{\mathrm{T}}\boldsymbol{V}^{\uparrow}_{\boldsymbol{s}_i}\boldsymbol{e}. \end{aligned} \quad (26)$$

Because both messages $m^{\mathrm{BP}}_{f_{G_i}\to x_i}(x_i)$ and $n_{x_i\to f_{G_i}}(x_i)$ in (9) are Gaussian, we obtain from (5)

$$m_{f_{G_i}\to x_i}(x_i) = m^{\mathrm{BP}}_{f_{G_i}\to x_i}(x_i). \quad (27)$$

5) Decoding: Decoding is performed by using the BCJR algorithm, which is an instance of BP [10]. After completion of the forward/backward processing the BCJR decoder returns the messages $m^{\mathrm{BP}}_{f_{M_i}\to x_i}(x_i)$, $\forall i$. We remark that any other code that can be decoded using a BP-based algorithm, e.g. a turbo- or LDPC code, could be used instead within the proposed BP-EP framework.

B. Scheduling of the messages

After initializing $m_{f_{M_i}\to x_i}(x_i)$, $\forall i$, the messages $m_{f_{G_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i)$ and $n_{\boldsymbol{s}_i\to f_{G_{i+1}}}(\boldsymbol{s}_i)$, with i ranging from 1 to $N+L-1$, are calculated in the downward recursion using (12) and (15) respectively. Likewise, $m_{f_{G_{i+1}}\to\boldsymbol{s}_i}(\boldsymbol{s}_i)$ and $n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i)$, with i ranging from $N+L-1$ to 1, are obtained from (18) and (21), respectively, in the upward recursion[3]. Equations (24) and (27) are used to get the messages $m_{f_{G_i}\to x_i}(x_i)$, $\forall i$, which are then passed to the BCJR decoder. The decoder outputs the messages $m^{\mathrm{BP}}_{f_{M_i}\to x_i}(x_i)$, $\forall i$, and finally $m_{f_{M_i}\to x_i}(x_i)$, $\forall i$, are updated via (9).

[3]Note that these recursions coincide with those of a Kalman smoother [7].

C. Reduction of complexity

Since they are performed in the update of each symbol, the two matrix inversions in (25) and (26) make up a significant part of the computational complexity of the BP-EP-based algorithm. To reduce the complexity, the approach proposed in [5] can be adopted. We calculate the belief of variable $\boldsymbol{s}_i$, i.e. $b(\boldsymbol{s}_i) \propto m_{f_{G_i}\to\boldsymbol{s}_i}(\boldsymbol{s}_i)\, n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i) \propto \mathcal{N}(\boldsymbol{s}_i;\boldsymbol{m}_i,\boldsymbol{V}_i)$, where

$$\begin{aligned}\boldsymbol{V}_i &= (\boldsymbol{V}^{\Downarrow\ -1}_{\boldsymbol{s}_i} + \boldsymbol{V}^{\uparrow\ -1}_{\boldsymbol{s}_i})^{-1}\\ \boldsymbol{m}_i &= \boldsymbol{V}_i(\boldsymbol{V}^{\Downarrow\ -1}_{\boldsymbol{s}_i}\boldsymbol{m}^{\Downarrow}_{\boldsymbol{s}_i} + \boldsymbol{V}^{\uparrow\ -1}_{\boldsymbol{s}_i}\boldsymbol{m}^{\uparrow}_{\boldsymbol{s}_i}).\end{aligned} \tag{28}$$

According to the deterministic relationship given in (3), the messages from factor node $f_{G_{i-l}}$ to variable nodes x_{i-l}, $l = 0, 1, \ldots, L-1$, are obtained as

$$\begin{aligned}\widetilde{m}^{\text{BP}}_{f_{G_{i-l}}\to x_{i-l}}(x_{i-l}) &\propto \frac{\int b(\boldsymbol{s}_i)\delta\left(s_{i,L-l} - x_{i-l}\right) d\boldsymbol{s}_i}{n_{x_{i-l}\to f_{G_{i-l}}}(x_{i-l})}\\ &\propto \exp\left\{-\frac{(x_{i-l} - \vec{m}_{x_{i-l}})^2}{2\vec{v}_{x_{i-l}}}\right\}\end{aligned} \tag{29}$$

where

$$\vec{v}_{x_{i-l}} = (V^{-1}_{i,L-l} - \overleftarrow{m}^{-1}_{x_{i-l}})^{-1} \tag{30}$$

$$\vec{m}_{x_{i-l}} = \vec{v}_{x_{i-l}}(V^{-1}_{i,L-l}m_{i,L-l} - \overleftarrow{v}^{-1}_{x_{i-l}}\overleftarrow{m}_{x_{i-l}}) \tag{31}$$

with $s_{i,L-l}$ and $m_{i,L-l}$ representing the $(L-l)$th element of vector $\boldsymbol{s}_i$ and $\boldsymbol{m}_i$ respectively, and $V_{i,L-l}$ denoting the $(L-l)$th diagonal entry of the matrix $\boldsymbol{V}_i$. Using this, only two matrix inversions are needed in the update of each block of L symbols [5]. Replacing the messages in (24) by those in (29) reduces the complexity order from $\mathcal{O}(L^3)$ to $\mathcal{O}(L^2)$. The proof of the equivalence of these messages is provided in Appendix A.

V. Simulation Results

We evaluate the performance of the communication system described in Section II by means of Monte Carlo simulations. Two different lengths of information bit vectors are considered: $K = 32768$ (long) and $K = 8192$ (short). The information bits are coded using a $1/2$ rate convolutional code $(23, 35)_8$. The vector of channel weights is set to $\boldsymbol{h} = [0.227\ 0.460\ 0.668\ 0.460\ 0.227]^{\mathrm{T}}$, which corresponds to a severe time-dispersive (5-tap) channel [11].

Fig. 2 and Fig. 3 depict the performance of the investigated algorithms: the proposed algorithm (BP-EP), the algorithm presented in [5] (GABP), the algorithm implementing MAP equalization (BP) (reproduced from Fig. 5 in [5]), and a receiver operating in an ISI-free channel (AWGN). In Fig. 2, the BER performance after 30 receiver iterations is shown when the SNR ranges from 4dB to 6dB. We observe that BP-EP significantly outperforms GABP. It also performs close to BP, the loss expressed in terms of the SNR value where the threshold effect occurs being about 0.3 dB. In Fig. 3, the BER performance at 5.5dB SNR of BP-EP and GABP is depicted as a function of the iteration index. We observe that BP-EP converges much faster and is less sensitive to shorter codeword lengths than GABP.

Both BP-EP and GABP receivers exhibit the same complexity order per symbol. They differ only in their respective equalization parts, both having $\mathcal{O}(L^2)$ order of complexity per symbol. The former algorithm approximates the messages from f_{M_i} to x_i based on the messages passed by both the decoder and the equalizer, while the latter only makes use of the messages passed by the decoder for doing this. The observed superior performance indicates that the BP-EP approximation is better.

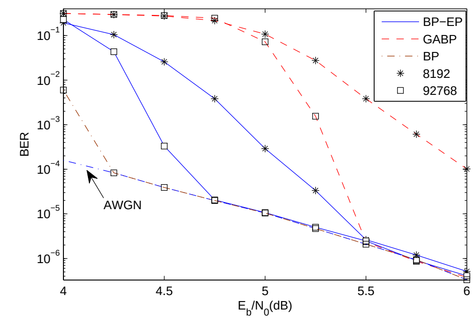

Fig. 2. BER performance versus E_b/N_0 of the investigated receivers.

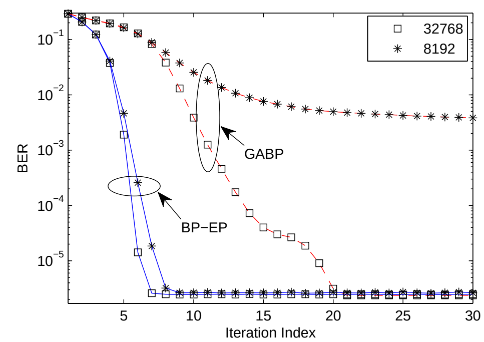

Fig. 3. BER performance versus iteration index of the investigated receivers.

Appendix A
Proof of the Equivalence Between (24) and (29)

The proof is by induction. Thus, we merely need to show the equivalence for $l = 0$ and $l = 1$.

For $l = 0$ we have according to (29)

$$\begin{aligned}&\widetilde{m}^{\text{BP}}_{f_{G_i}\to x_i}(x_i) \propto \textstyle\int b(\boldsymbol{s}_i)\delta\left(s_{i,L} - x_i\right) d\boldsymbol{s}_i / n_{x_i\to f_{G_i}}(x_i)\\ &\propto \textstyle\int n_{\boldsymbol{s}_{i-1}\to f_{G_i}}(\boldsymbol{s}_{i-1})\, n_{x_i\to f_{G_i}}(x_i) f_{G_i}(\boldsymbol{s}_i, \boldsymbol{s}_{i-1}, x_i)\, d\boldsymbol{s}_{i-1}dx_i\\ &\qquad\qquad \times n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i)\,\delta\left(s_{i,L} - x_i\right) d\boldsymbol{s}_i / n_{x_i\to f_{G_i}}(x_i)\\ &= \textstyle\int n_{\boldsymbol{s}_{i-1}\to f_{G_i}}(\boldsymbol{s}_{i-1})\, n_{\boldsymbol{s}_i\to f_{G_i}}(\boldsymbol{s}_i)\, f_{G_i}(\boldsymbol{s}_i, \boldsymbol{s}_{i-1}, x_i)\, d\boldsymbol{s}_{i-1}d\boldsymbol{s}_i\\ &= m^{\text{BP}}_{f_{G_i}\to x_i}(x_i).\end{aligned} \tag{32}$$

For $l = 1$ we first obtain from the BP rule (4)

$$\begin{aligned}&\textstyle\int b(\boldsymbol{s}_i)\delta\left(s_{i,L-1} - x_{i-1}\right) d\boldsymbol{s}_i\\ &\qquad \propto \textstyle\int b(\boldsymbol{s}_{i-1})\delta\left(s_{i-1,L} - x_{i-1}\right) d\boldsymbol{s}_{i-1}.\end{aligned} \tag{33}$$

Then, using (32) and (33) yields

$$\begin{aligned}&\widetilde{m}^{\text{BP}}_{f_{G_{i-1}}\to x_{i-1}}(x_{i-1}) \propto \frac{\int b(\boldsymbol{s}_i)\delta\left(s_{i,L-1} - x_{i-1}\right) d\boldsymbol{s}_i}{n_{x_{i-1}\to f_{G_{i-1}}}(x_{i-1})}\\ &\propto \frac{\int b(\boldsymbol{s}_{i-1})\delta\left(s_{i-1,L} - x_{i-1}\right) d\boldsymbol{s}_{i-1}}{n_{x_{i-1}\to f_{G_{i-1}}}(x_{i-1})} \propto m^{\text{BP}}_{f_{G_{i-1}}\to x_{i-1}}(x_{i-1}).\end{aligned}$$

REFERENCES

[1] R. Koetter, A. C. Singer, and M. Tücher, "Turbo equalization," *IEEE Signal Processing Mag.*, vol. 21, pp. 67–80, Jan. 2004.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE International Conference on Communications, (ICC'93)*, May 1993, pp. 1064–1070.

[3] R. McEliece, D. MacKay, and J. Cheng, "Turbo decoding as an instance of Pearl's "belief propagation" algorithm," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb. 1998.

[4] G. Colavolpe and G. Germi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 818–825, May 2005.

[5] Q. Guo and L. Ping, "LMMSE turbo equalization based on factor graphs," *IEEE J. Select. Areas Commun.*, vol. 26, no. 2, pp. 311–319, Feb. 2008.

[6] X. Wang and H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, Jul. 1999.

[7] J. Hu, H.-A. Loeliger, J. Dauwels, and F. Kschischang, "A general computation rule for lossy summaries/messages with examples from equalization," in *Proc. 44th Allerton Conf. Communication, Control, and Computing*, Sep. 2006, pp. 27–29.

[8] T. Heskes, M. Opper, W. Wiegerinck, O. Winther, and O. Zoeter, "Approximate inference techniques with expectation constraints," *Journal of Statistical Mechanics – Theory and Experiment*, Sep. 2005.

[9] T. P. Minka, "Expectation propagation for approximate Bayesian inference," in *Proceedings of the Seventeenth Conf. on Uncertainty in Artificial Intelligence*, 2001, pp. 362–369.

[10] 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.

[11] J. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, Aug. 2000.