Message-Passing Algorithms for Channel Estimation and Decoding Using Approximate Inference

Mihai-A. Badiu‡∗, Gunvor E. Kirkelund*, Carles Navarro Manchón*, Erwin Riegler‡ and Bernard H. Fleury*

∗ Aalborg University, Denmark
‡ Technical University of Cluj-Napoca, Romania

—W e design iterative receiver schemes for a generic communication system by treating channel estimation and information decoding as an inference problem in graphical models. We introduce a recently proposed inference framework that combines belief propagation (BP) and the mean field (MF) approximation and includes these algorithms as special cases. We also show that the expectation propagation and expectation maximization (EM) algorithms can be embedded in the BP-MF framework with slight modifications. By applying the considered inference algorithms to our probabilistic model, we derive four different message-passing receiver schemes. Our numerical evaluation in a wireless scenario demonstrates that the receiver based on the BP-MF framework and its variant based on BP-EM yield the best compromise between performance, computational complexity and numerical stability among all candidate algorithms.

I. INTRODUCTION

The design of advanced receiver algorithms is crucial to meet the stringent requirements of modern communication systems. Motivated by the successful application of the “turbo” principle in the decoding of channel codes, a large number of works have been devoted to the design of turbo receivers (see [1] and the references therein). While in many of these works the receiver modules are individually designed and heuristically interconnected to exchange soft values, iterative receiver algorithms can be rigorously designed and better understood as instances of message-passing inference techniques (e.g., see [2]).

In this context, variational Bayesian inference in probabilistic models [3] have proven to be a very useful tool to design receivers where tasks like channel estimation, detection and decoding are jointly derived. Among the variational techniques, belief propagation (BP) [4], [5] has found the most widespread use. Originally applied to the decoding of channel codes, BP has been shown to be especially efficient in discrete probabilistic models. An alternative to BP is the mean field (MF) approximation and its message-passing counterpart, usually referred to as variational message-passing [6]. MF inference has been successfully applied to continuous probabilistic models involving probability density functions (pdfs) belonging to an exponential family, in which BP suffers from numerical intractability. Other notable examples of general-purpose inference techniques are expectation-maximization (EM) [7] and expectation propagation (EP) [8]. EM is a special case of MF, where the approximate pdfs – referred to as beliefs – are Dirac delta functions; EP can be seen as an approximation of BP where some beliefs are approximated by pdfs in a specific exponential family. Some attempts to find a unified framework encompassing all these techniques include the α-divergence interpretation in [9] and the region-based free energy approximations in [10]. Following the latter approach, a novel hybrid message-passing inference framework combining BP and the MF approximation was recently proposed in [11].

In this paper, we investigate the design of receivers that perform joint channel estimation and data decoding in a generic communication system. For this purpose, we capitalize on the combined inference framework [11], which provides some degree of freedom in the choice of the parts of the factor graph in which either BP or MF is applied. We show that this framework can be modified to naturally embed EP, EM and BP with Gaussian approximation of some messages. Then, we apply these hybrid inference techniques to the underlying probabilistic model of the system and obtain four receiver algorithms, whose performance we assess by simulating a wireless system.

Notation: we denote by |I| the cardinality of a finite set I; the relative complement of {i} in I is written as I \ {i}; the set \{i ∈ N | 1 ≤ i ≤ n\} is denoted by [1 : n]. Boldface lowercase and uppercase letters are used to represent vectors and matrices, respectively; superscripts (·)T and (·)H denote transposition and Hermitian transposition, respectively. The Hadamard product of two vectors is denoted by ⊙. For a vector x = (x_i | i ∈ I)T, we write x_i = (x_j | j ∈ I \ {i})T; for a matrix A ∈ Cm×n, [A]_{i,j} denotes its (i,j)th entry, [A]_{i,j} denotes the matrix A with the ith row and jth column deleted, [A]_{i,j} denotes the column vector ([A]_{i,j} | k ∈ [1 : n])T, and [A]_{i,j} denotes the row vector ([A]_{i,j} | k ∈ [1 : n]). The pdf of a multivariate complex Gaussian distribution with mean µ and covariance matrix Σ is denoted by CN(·; µ, Σ). We write f(x) ∝ g(x) when f(x) = cg(x) for some positive constant c. We denote by G[·] the approximation of the pdf in the argument with a Gaussian pdf with the same mean and covariance matrix. The Dirac delta function is denoted by δ(·).

II. MESSAGE-PASSING INFERENCE ALGORITHMS

We begin by concisely describing the unified message-passing algorithm that combines the BP and MF approaches (refer to [11]). Then, we briefly show how other widespread inference algorithms can be obtained as particular instances or slight modifications of the unified framework.

Let p(z) be an arbitrary pdf of a random vector z ∈ I T which factorizes as

\[ p(z) = \prod_{a ∈ A} f_a(z_a) = \prod_{a ∈ A_{MF}} f_a(z_a) \prod_{c ∈ A_{BP}} f_c(z_c) \]  

(1)

where z_a is the vector of all variables z_i that are arguments of the function f_a. We have grouped the factors into two sets that partition A: \( A_{MF} \cap A_{BP} = \emptyset \) and \( A_{MF} \cup A_{BP} = A \).
factorization in (1) can be visualized in a factor graph [4]
representation. We define \( \mathcal{N}(a) \subseteq \mathcal{I} \) to be the set of indices
of all variables \( z_i \) that are arguments of function \( f_a \); similarly,
\( \mathcal{N}(i) \subseteq \mathcal{A} \) denotes the set of indices of all functions \( f_a \)
that depend on \( z_i \). The parts of the graph that correspond to \( \prod_{a \in \mathcal{A}_{BP}} f_a(z_a) \) and to \( \prod_{a \in \mathcal{A}_{MF}} f_a(z_a) \) are referred to as “BP
part” and “MF part”, respectively. We denote the variable nodes in the
BP part by \( \mathcal{I}_{BP} \equiv \bigcup_{a \in \mathcal{A}_{BP}} \mathcal{N}(a) \) and those in the MF
part by \( \mathcal{I}_{MF} \equiv \bigcup_{a \in \mathcal{A}_{MF}} \mathcal{N}(a) \).

The combined BP-MF inference algorithm approximates the
marginals \( p(z_i) = \int p(z)dz_i, i \in \mathcal{I} \) by auxiliary pdfs \( b_i(z_i) \)
called beliefs. They are computed as [11]

\[
b_i(z_i) = \omega_i \prod_{a \in \mathcal{A}_{BP} \cap \mathcal{N}(i)} m_{a \rightarrow i}^{BP}(z_i) \prod_{a \in \mathcal{A}_{MF} \cap \mathcal{N}(i)} m_{a \rightarrow i}^{MF}(z_i) \tag{2}
\]
with

\[
m_{a \rightarrow i}^{BP}(z_i) = \omega_a \int \prod_{j \in \mathcal{N}(a) \setminus i} dz_j n_{j \rightarrow a}(z_j) f_a(z_a), \quad \forall a \in \mathcal{A}_{BP}, i \in \mathcal{N}(a)
\]

\[
m_{a \rightarrow i}^{MF}(z_i) = \exp \left( \int \prod_{j \in \mathcal{N}(a) \setminus i} dz_j n_{j \rightarrow a}(z_j) \ln f_a(z_a) \right), \quad \forall a \in \mathcal{A}_{MF}, i \in \mathcal{N}(a)
\]

\[
n_{i \rightarrow a}(z_i) = \omega_i \prod_{c \in \mathcal{A}_{BP} \cap \mathcal{N}(i) \setminus a} m_{c \rightarrow i}^{BP}(z_i) \prod_{c \in \mathcal{A}_{MF} \cap \mathcal{N}(i)} m_{c \rightarrow i}^{MF}(z_i), \quad \forall i \in \mathcal{N}(a), a \in \mathcal{A}, \tag{3}
\]

where \( \omega_i \) and \( \omega_a \) are constants that ensure normalized beliefs.

Belief propagation is obtained as a particular case of BP-MF
by setting \( \mathcal{A}_{MF} = \emptyset \), since in this case the expressions in (3)
reduce to the BP message computations. Similarly, mean field is
an instance of BP-MF when \( \mathcal{A}_{BP} = \emptyset \).

Expectation propagation is very similar to BP, the main
difference being that it constrains the beliefs of some variables
to be members of a specific exponential family. Assuming Gaussian
approximations of the beliefs, EP can also be integrated in the
BP-MF framework by modifying the messages

\[
m_{a \rightarrow i}^{EP}(z_i) \propto \frac{1}{n_{i \rightarrow a}(z_i)} \mathbb{E} \left[ n_{i \rightarrow a}(z_i) m_{a \rightarrow i}^{BP}(z_i) \right], \tag{4}
\]
for all \( i \in \mathcal{I}_{BP} \subseteq \mathcal{I}_{BP}, a \in \mathcal{N}(i) \cap \mathcal{A}_{BP} \).

The expectation-maximization algorithm is a special case of
MF when the beliefs of some variables are constrained to be
Dirac delta functions [11]. Again, we include this approximation
in the BP-MF framework. This leads to \( n_{i \rightarrow a}(z_i) = \delta(z_i - \bar{z}_i) \)
for all \( i \in \mathcal{I}_{EM} \subseteq \mathcal{I}_{MF} \) and \( a \in \mathcal{N}(i) \cap \mathcal{A}_{MF} \), where \( \bar{z}_i \) maximizes the unconstrained belief (2). We refer to this modified algorithm
as BP-EM.

### III. Probabilistic System Model

In this section, we present the signal model of our inference
problem and its graphical representation. These will establish
the baseline for the derivation of message-passing receivers.

We analyze a system consisting of one transmitter and one receiver. A message represented by a vector \( \mathbf{u} = (u_k | k \in [1:K])^T \in \{0,1\}^K \) of information bits is conveyed
by sending \( N \) data and \( M \) pilot channel symbols having the
sets of indices \( \mathcal{D} \subseteq [1:M+N] \) and \( \mathcal{P} \subseteq [1:M+N] \),
respectively, such that \( \mathcal{D} \cup \mathcal{P} = [1:M+N] \) and \( \mathcal{D} \cap \mathcal{P} = \emptyset \).

Specifically, vector \( \mathbf{u} \) is encoded and interleaved using a rate
\( R = K/(NL) \) channel code and a random interleaver into the
vector \( \mathbf{c} = (c_i^T | c_i \in \{0,1\}^L, n \in [1:N])^T \) of length \( NL \).
For each \( n \in [1:N] \), the subvector \( \mathbf{c}_n = (c_n^1, \ldots, c_n^L)^T \) is
mapped to a data symbol \( x_{ni} \in \mathcal{S}_D \) with \( i_n \in \mathcal{D} \), where \( \mathcal{S}_D \) is a
discrete complex modulation alphabet of size \( 2^L \). Symbols
\( x_D = (x_{i} | i \in \mathcal{D})^T \) are multiplexed with pilot symbols
\( x_P = (x_{j} | j \in \mathcal{P})^T \), which are randomly selected from a QPSK
modulation alphabet. Finally, the aggregate vector of channel
symbols \( \mathbf{x} = (x_i | i \in \mathcal{D} \cup \mathcal{P})^T \) is sent through a channel with
the following input-output relationship:

\[
y = h \odot x + w. \tag{5}
\]
The vector \( \mathbf{y} = (y_i | i \in [1:M+N]^T \) contains the received
signal samples, \( h = (h_i | i \in [1:M+N])^T \) is the vector of channel
coefficients, and \( w = (w_i | i \in [1:M+N])^T \) contains the samples of additive noise and has the pdf \( p(w) = CN(\mathbf{w};0,\gamma^{-1}I_{M+N}) \) for some positive component precision \( \gamma \).

Note that (5) can model any channel with a multiplicative effect
that is not affected by inter-symbol interference, e.g., a
time-varying frequency-flat channel or the equivalent channel in
the frequency domain in a multicarrier system.

Based on the above signal model, we can state the probabilistic
model which captures the dependencies between the system
variables. The pdf of the collection of observed and unknown
variables factorizes as

\[
p(\mathbf{y}, \mathbf{x}_D, \mathbf{c}, \mathbf{u}) = f_{h}(\mathbf{h}) \prod_{i \in \mathcal{D}} f_{d_i}(h_i, x_i) \prod_{j \in \mathcal{P}} f_{p_j}(h_j) \times \prod_{n \in [1:N]} f_{M_n}(x_{i_n}, c_n) f_{c}(\mathbf{c}, \mathbf{u}) \prod_{k \in [1:K]} f_{\mu_k}(u_k), \tag{6}
\]
where \( f_{d_i}(h_i, x_i) \triangleq p(y_i|h_i, x_i) \) and \( f_{p_j}(h_j) \triangleq p(y_j|h_j) \) incorporate the observations in \( \mathbf{y} \) and are given by

\[
f_{d_i}(h_i, x_i) = \text{CN}(h_i x_i; y_i, \gamma^{-1}), \quad \forall i \in \mathcal{D},
\]

\[
f_{p_j}(h_j) = \text{CN}(h_j x_j; y_j, \gamma^{-1}), \quad \forall j \in \mathcal{P}, \tag{7}
\]

\[
f_{h}(\mathbf{h}) \triangleq p(\mathbf{h}) \text{ is the prior pdf of the vector of channel coefficients for which we set}
\]

\[
f_{h}(\mathbf{h}) = \text{CN}(\mathbf{h}; \mu^h, \Sigma^h), \tag{8}
\]

\[
f_{M_n}(x_{i_n}, c_n) \triangleq p(x_{i_n} | c_n) \text{ stand for the modulation mapping,}
\]

\[
f_{c}(\mathbf{c}, \mathbf{u}) \triangleq p(\mathbf{c} | \mathbf{u}) \text{ accounts for the coding and interleaving operations and}
\]

\[
f_{\mu_k}(u_k) \triangleq p(u_k) \text{ is the prior pmf of the } k\text{th information bit. To obtain (6), we used the fact that } \mathbf{y} \text{ is conditionally independent of } \mathbf{c} \text{ and } \mathbf{u} \text{ given } \mathbf{x}_D, \text{ } \mathbf{c} \text{ is independent of}
\]

\[
\text{and } \mathbf{u}, \text{ the noise samples } w_i \text{ are i.i.d., and each data symbol } x_{i_n} \text{ is conditionally independent of all the other symbols}
\]

\[
given c_n. The factorization in (6) can be visualized in the factor
graph depicted in Fig. 1. The graph of the code and interleaver
is not explicitly given, its structure being captured by } f_c.

### IV. Message-Passing Receiver Schemes

In this section, we derive iterative receiver schemes by
applying different inference algorithms to the factor graph in
obtain the messages

\[ \text{Fig. 1. The receiver has to infer the beliefs of the information bits.} \]

Since the intractability of the messages occurs due to the structure of the channel code and interleaver.

We set \( \mathcal{A} \) and \( \mathcal{I} \) (defined in Section II for a general probabilistic model) to be the sets of all factors and variables, respectively, contained in our probabilistic model. Next, we show that the BP algorithm resulting from setting \( \mathcal{A}_{\text{BF}} = \emptyset \) yields messages of an intractable complexity. Assume that by running BP in the part of the graph containing the modulation and code constraints we obtain the messages

\[ m_{f_{\mathcal{m}} \rightarrow x_{i_n}}^{\text{BP}}(x_{i_n}) \propto \sum_{s \in \mathcal{S}_0} \beta_{i_n}(s) \delta(x_{i_n} - s), \]

with \( i_n \in \mathcal{D}, \forall n \in [1 : N] \), where \( \beta_{i_n}(s) \) represent extrinsic information on symbol \( x_{i_n} \). These messages are further passed as \( n_{x_{i_n} \rightarrow f_{\mathcal{m}}}(x_{i_n}) = m_{f_{\mathcal{m}} \rightarrow x_{i_n}}^{\text{BP}}(x_{i_n}) \). Then, for each \( i \in \mathcal{D} \), compute the message

\[ m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \propto \int f_{D_i}(h_i, x_i) n_{x_i \rightarrow f_{\mathcal{m}}}(x_i) \, dx_i \]

\[ \propto \sum_{s \in \mathcal{S}_0} \beta_{i_n}(s) \frac{\gamma_i y_i s^*}{|s|^2} \left( \frac{1}{\gamma_i |s|^2} \right), \]

for all \( j \in \mathcal{P} \) set

\[ m_{f_{\mathcal{m}} \rightarrow h_j}(h_j) \propto \frac{f_{\mathcal{m}}(h_j)}{f_{\mathcal{m}}(h_j)} \propto \left( \frac{y_j x_j^*}{|x_j|^2} \right)^2 \left( \frac{1}{\gamma |x_j|^2} \right). \]

Note that the message in (11) is proportional to a mixture of Gaussian pdfs with \( |\mathcal{S}_0| = 2^L \) components. Then, after setting \( n_{h_i \rightarrow f_{\mathcal{m}}}(h_i) = m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \) for all \( i \in \mathcal{D} \) and \( n_{h_i \rightarrow f_{\mathcal{m}}}(h_i) = m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \) for all \( i \in \mathcal{P} \), the message from \( f_{\mathcal{m}} \) to \( h_i \) reads

\[ m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \propto \int f_{D_i}(h_i, x_i) n_{x_i \rightarrow f_{\mathcal{m}}}(x_i) \, dh_j. \]

Using (9), (11) and (12), the message in (13) becomes a Gaussian mixture with \( 2^L(N-1) \) and \( 2^L N \) components for \( i \in \mathcal{D} \) and \( i \in \mathcal{P} \), respectively. Clearly, the computation of such messages is intractable and one has to use approximations.

A. Algorithm based on BP combined with Gaussian approximation

Since the intractability of the messages occurs due to the Gaussian mixture in (11), we approximate those messages as proposed in [12], i.e., for each \( i \in \mathcal{D} \) we set

\[ m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \propto \mathcal{G}(h_i; \mu_{h_i}, \sigma_{h_i}^2) \]

with

\[ \mu_{h_i} = \alpha_i(s) \frac{y_i s^*}{|s|^2}, \]

\[ \sigma_{h_i}^2 = \alpha_i(s) \left( \frac{\mu_{h_i}^2}{|s|^2} + \frac{1}{|\gamma| |s|^2} \right) - |\mu_{h_i}|^2, \]

In (15), we have defined the normalized amplitudes of the Gaussian mixture \( \alpha_i(s) = \beta_i(s)/(|h_i| |s|^2) \), where the constant \( \kappa_i \) ensures \( \sum_{s \in \mathcal{S}_0} \alpha_i(s) = 1 \). We also denote the mean and variance of the pdf in (12) by \( \mu_{h_i} \) and \( \sigma_{h_i}^2 \), \( j \in \mathcal{P} \), and we define the vector \( \mu_{h_i} = (\mu_{h_i} \ldots \mu_{h_i})^T \) and the matrix \( \Sigma_{h_i} \) with entries \( \Sigma_{h_i,j} = \sigma_{h_i}^2 \) if \( i = j \) and zero otherwise, for all \( i, j \in [1 : M + N] \).

Now, using (9) and (14), the message in (13) becomes

\[ m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \propto \int \mathcal{G}(h_i; \mu_{h_i}, \Sigma_{h_i}) \, dh_i \]

\[ \propto \mathcal{G}(h_i; \mu_{h_i}, \sigma_{h_i}^2), \]

with

\[ \mu_{h_i} = \mu_{h_i} + \left( \Sigma_{h_i,i} + \Sigma_{h_i,i}^{-1} (\mu_{h_i} - \mu_{h_i}) \right)^{-1}, \]

\[ \sigma_{h_i}^2 = \left( \Sigma_{h_i,i}^{-1} - \left( \Sigma_{h_i,i}^{-1} + \Sigma_{h_i,i}^{-1} \right)^{-1} \right) \Sigma_{h_i,i}. \]

These messages are further passed as extrinsic values, i.e., \( n_{h_i \rightarrow f_{\mathcal{m}}}(h_i) = m_{f_{\mathcal{m}} \rightarrow h_i}^{\text{BP}}(h_i) \). For each \( i \in \mathcal{D} \), the following message is then computed:

\[ m_{f_{\mathcal{m}} \rightarrow x_i}(x_i) \propto \int f_{D_i}(h_i, x_i) n_{h_i \rightarrow f_{\mathcal{m}}}(h_i) \, dh_i \]

\[ \propto \frac{1}{\gamma - 1 + \sigma_{h_i}^2 |x_i|^2} \exp \left( -\frac{|y_i - \mu_{h_i} x_i|^2}{\gamma - 1 + \sigma_{h_i}^2 |x_i|^2} \right). \]

After passing the extrinsic messages \( n_{x_{i_n} \rightarrow f_{\mathcal{m}}}(x_{i_n}) = m_{f_{\mathcal{m}} \rightarrow x_{i_n}}^{\text{BP}}(x_{i_n}) \), we apply the BP update rule to compute the probabilities of the coded and interleaved bits (which is equivalent to MAP demapping), followed by BP decoding to obtain the beliefs of the information bits.

B. Algorithm based on expectation propagation

We set \( \mathcal{A}_{\text{EP}} = \emptyset \) and \( \mathcal{I}_{\text{EP}} = \{ h_i, i \in \mathcal{D} \} \). The message \( m_{f_{\mathcal{m}} \rightarrow h_i}(h_i) \) computed with (4) is proportional to a Gaussian pdf; consequently, the EP rule for \( m_{f_{\mathcal{m}} \rightarrow h_i}(h_i) \) reduces to the BP rule and outputs a Gaussian pdf as in (16), since the operator \( \mathcal{G}[\cdot] \) is an identity operator for Gaussian arguments.

Specifically, using (3), (4), and then (11), (16), we have

\[ h_i = \mathcal{G} \left( n_{h_i \rightarrow f_{\mathcal{m}}}(h_i) m_{f_{\mathcal{m}} \rightarrow h_i}(h_i) \right) = \mathcal{G}(h_i; \mu_{h_i}, \sigma_{h_i}^2), \]

for each \( i \in \mathcal{D} \), where

\[ \mu_{h_i} = \alpha_i(s) \frac{\sigma_{h_i}^2 \mu_{h_i} + \gamma y_i s^*}{\sigma_{h_i}^2 + |\gamma| |s|^2}, \]

\[ \sigma_{h_i}^2 = \alpha_i(s) \left( \frac{\sigma_{h_i}^2 \mu_{h_i} + \gamma y_i s^*}{\sigma_{h_i}^2 + |\gamma| |s|^2} \right)^2 - |\mu_{h_i}|^2. \]
with
\[ \phi_i(s) = \frac{\beta_i(s)}{\sum_{s' \in S} \beta_i(s')} \text{CN} \left( y_i; \mu_{h_i,c} s, \gamma^{-1} + \sigma_{h_i,c}^2 |s|^2 \right) \]
and \( \mu_{h_i,c}, \sigma_{h_i,c}^2 \) as in (17). Using (4) again, we obtain

\[ m_{f_{D_i} \rightarrow h_i}^{\text{EP}}(h_i) \propto \frac{\text{CN} \left( h_i; \mu_{h_i, o}, \sigma_{h_i, o}^2 \right)}{\text{CN} \left( h_i; \mu_{h_i, c}, \sigma_{h_i, c}^2 \right)} \]

with
\[ \sigma_{h_{i, o}}^{-2} = \sigma_{h_{i, c}}^{-2} - \sigma_{h_{i, o}}^{-2}, \]
\[ \mu_{h_{i, o}} = \sigma_{h_{i, c}}^{-2} \left( \sigma_{h_{i, c}}^{-2} - \mu_{h_{i, c}} \right). \] (18)

Unlike (15) in BP with Gaussian approximation, the values of \( \mu_{h_{i, o}} \) and \( \sigma_{h_{i, o}}^2 \) depend on all \( \mu_{h_{j, o}} \) and \( \sigma_{h_{j, o}}^2 \), \( j \in D \), \( j \neq i \), through (17). The parameters of \( m_{f_{D_i} \rightarrow h_i}^{\text{EP}}(h_i) \) are updated using (17) but with \( \mu_{h_{i, o}} \) and \( \sigma_{h_{i, o}}^2 \) computed as above. Note that all messages that depend on the channel coefficients need to be updated in a sequential manner. The rest of the messages are computed as in Section IV-A.

C. Algorithm based on the combined BP-MF framework

The factor graph is split into the MF and BP parts by setting \( A_{\text{MF}} = \{ f_D \mid i \in D \} \) and \( A_{\text{BP}} = A \setminus A_{\text{MF}} \). Such a splitting yields tractable and simple messages, takes advantage of the fact that BP works well with hard constraints and best exploits the correlation between the channel coefficients for the graphical representation in Fig. 1.

Assuming we have obtained the messages \( m_{x_i \rightarrow f_{D_i}}(x_i) \) (their expression will be given later), we can compute

\[ m_{f_{D_i} \rightarrow h_i}^{\text{MF}}(h_i) \propto \exp \left( \int m_{x_i \rightarrow f_{D_i}}(x_i) \ln f_{D_i}(h_i, x_i) \, dx_i \right) \]
\[ \propto \text{CN} \left( h_i; \mu_{h_i, o}, \sigma_{h_i, o}^2 \right), \]

where
\[ \mu_{h_i, o} = \frac{y_i |h_i|^2}{\gamma \left( \sigma_{h_i}^2 + |x_i|^2 \right)}, \quad \sigma_{h_i, o}^2 = \frac{1}{\gamma \left( \sigma_{h_i}^2 + |x_i|^2 \right)}, \]

with the definition \( \mu_{x_i} \equiv \int m_{x_i \rightarrow f_{D_i}}(x_i) x_i \, dx_i \) and \( \sigma_{x_i}^2 \equiv \int m_{x_i \rightarrow f_{D_i}}(x_i) |x_i|^2 \, dx_i \).

The messages \( m_{h_i \rightarrow f_{D_i}}(h_i) = m_{f_{D_i} \rightarrow h_i}^{\text{MF}}(h_i) \) are sent to the BP part and hence are extrinsic values. When computing \( m_{h_i \rightarrow f_{D_i}}^{\text{BP}}(h_i) \) we get the same expression as (16), with the parameters (17). Unlike the previous algorithms, the following messages are beliefs, i.e., \( a \text{ posteriori} \) probabilities (APP):

\[ n_{h_i \rightarrow f_{D_i}}(h_i) = \omega_{h_i} m_{f_{D_i} \rightarrow h_i}^{\text{BP}}(h_i) m_{f_{D_i} \rightarrow h_i}^{\text{MF}}(h_i) \]
\[ = \text{CN} \left( h_i; \mu_{h_i, c}, \sigma_{h_i, c}^2 \right), \quad \forall i \in D, \]

with
\[ \mu_{h_i} = \left( \sigma_{h_{i, o}}^{-2} + \sigma_{h_{i, c}}^{-2} \right)^{-1} \left( \sigma_{h_{i, o}}^{-2} \mu_{h_{i, o}} + \sigma_{h_{i, c}}^{-2} \mu_{h_{i, c}} \right), \]
\[ \sigma_{h_{i, o}}^{-2} = \sigma_{h_{i, c}}^{-2} + \sigma_{h_{i, c}}^{-2}. \] (19)

Then, for all \( i \in D \), we compute

\[ m_{f_{D_i} \rightarrow x_i}(x_i) \propto \exp \left( \int n_{h_i \rightarrow f_{D_i}}(h_i) \ln f_{D_i}(h_i, x_i) \, dh_i \right) \]
\[ \propto \text{CN} \left( x_i; \frac{y_i |h_i|^2}{\gamma \sigma_{h_i}^2 + |x_i|^2}, \frac{1}{\gamma |x_i|^2} \right) \] (20)

and we pass \( n_{x_i \rightarrow f_{D_i}}(x_i) = m_{f_{D_i} \rightarrow x_i}(x_i) \) to the modulation and coding part of the graph as extrinsic values, for all \( n \in [1 : N] \). After running BP, we obtain (10) and then pass the following APP values back to the MF part:

\[ n_{x_i \rightarrow f_{D_i}}(x_i) = \omega_{x_i} m_{f_{D_i} \rightarrow x_i}^{\text{BP}}(x_i) m_{f_{D_i} \rightarrow x_i}^{\text{MF}}(x_i). \]

D. Algorithm based on BP-EM

We now apply EM for channel estimation, so we constrain \( b_{h_i}(h_i) \) from the previous BP-MF scheme to be Dirac delta functions. The resulting messages are the same as in the previous subsection, except for \( n_{h_i \rightarrow f_{D_i}}(h_i) = \delta(h_i - \mu_{h_i}) \) with \( \mu_{h_i} \) computed as in (19). Note that this algorithm uses only point estimates of the channel weights; however, its complexity is basically still the same, since the computation of (19) actually includes the computation of the corresponding variance.

E. Scheduling of message computations

All algorithms employ the same message-passing scheduling: they start by sending messages \( m_{f_{D_i} \rightarrow h_i}(h_i) \) corresponding to pilots and by initializing \( m_{f_{D_i} \rightarrow h_i}(h_i) \propto \text{CN}(h_i; 0, \infty) \); messages (computed according to the corresponding algorithm) are passed up to the information bit variables – this completes the first iteration; each following iteration consists in passing messages up to the channel prior factor node and back; messages are passed back and forth until a predefined number of iterations is reached. All algorithms end by taking hard decisions on the beliefs of the information bits.

V. SIMULATION RESULTS

We consider a wireless OFDM system with the parameters given in Table I, and we evaluate by means of Monte Carlo simulations the bit error rate (BER) performance of the receiver algorithms derived in Section IV. We employ as a reference a scheme which has perfect channel state information (CSI), i.e., it has prior knowledge of the vector of channel coefficients \( \mathbf{h} \).

We encountered numerical problems with the EP-based scheme due to the instability of EP in general, so we used the heuristic approach [9] to damp the updates of the beliefs \( b_{h_i} \) with a step-size \( \epsilon = 0.5 \). Also, the EP-based scheme has higher computational complexity than the others due to its message definition – it requires multiplication of a Gaussian pdf with a mixture of Gaussian pdfs, the approximation \( G[\cdot] \) and division of Gaussian pdfs – and to the sequentiality of the message updates for the channel coefficients \( \mathbf{h} \).

Results in terms of BER versus signal-to-noise ratio (SNR) are given in Fig. 2, while the convergence of the BER with the number of iterations is illustrated in Fig. 3. The receivers based on EP, combined BP-MF and BP-EM exhibit similar

---

1Alternatively, the same level of exploitation of the correlation is obtained by representing the channel variables as a single vector variable \( \mathbf{h} \) and “moving” factor node \( f_D \) to the MF part [11].

2For the other receiver schemes, it can be shown that the parameters of all messages \( m_{f_{D_i} \rightarrow h_i}(h_i) \) with \( i \in D \cup P \) can be computed jointly and with a lower complexity.
We formulated the problem of joint channel estimation and decoding in a communication system as inference in a graphical model. To solve the inference problem, we resorted to a recently proposed message-passing framework that unifies the BP and MF algorithms and includes them as particular instances. Additionally, we illustrated how the combined framework can encompass the EP and EM inference algorithms.

Based on the inference techniques considered, we derived four receiver algorithms. Since BP is not suitable for the studied problem, as it leads to intractable messages, we applied its variant which employs Gaussian approximation of the computationally cumbersome messages instead. However, our results showed that it performs significantly worse than the other proposed schemes. Considering the BER results, the computational complexity and stability of these schemes, we conclude that the receiver based on the combined BP-MF framework and its BP-EM variant are the most effective receiver algorithms.

ACKNOWLEDGMENT

Six projects have supported this work: the Project SIDOC under contract no. POSDRU/88/1.5/S/60078; the Cooperative Research Project 4GMCT funded by Intel Mobile Communications, Agilent Technologies, Aalborg University and the Danish National Advanced Technology Foundation; the PhD Project “Iterative Information Processing for Wireless Receivers” funded by Renesas Mobile Corporation; the Project ICT-248894 WHERE2; the WWTF Grant ICT10-066; and the FWF Grant S10603-N13 within the National Research Network SISE.

REFERENCES


VI. CONCLUSIONS

We formulated the problem of joint channel estimation and decoding in a communication system as inference in a graphical model. To solve the inference problem, we resorted to a recently proposed message-passing framework that unifies the BP and MF algorithms and includes them as particular instances. Additionally, we illustrated how the combined framework can encompass the EP and EM inference algorithms.

Based on the inference techniques considered, we derived four receiver algorithms. Since BP is not suitable for the studied problem, as it leads to intractable messages, we applied its variant which employs Gaussian approximation of the computationally cumbersome messages instead. However, our results showed that it performs significantly worse than the other proposed schemes. Considering the BER results, the computational complexity and stability of these schemes, we conclude that the receiver based on the combined BP-MF framework and its BP-EM variant are the most effective receiver algorithms.