Iterative mitigation of intercell interference in cellular networks based on Gaussian belief propagation

Frederic Lehmann

Abstract

The mitigation of two-dimensional intersymbol interference has emerged recently as an important problem in the field of storage technology and wireless communications. Several solutions based on message passing have been proposed to perform equalization on two-dimensional channels with intersymbol interference. We present an algorithm based on Gaussian belief propagation. We show that the proposed algorithm has interesting features compared to the corresponding message passing solutions available in the literature. Firstly, the complexity of the proposed equalizer is independent of the size of the constellation used for modulation. Secondly, since the complexity is not exponential in the size of the state-space, the computational burden is reduced for channels with long memory. As an application, we consider decentralized intercell interference mitigation in uplink cellular networks, where base stations communicate with their neighbors to recover their own users' signals.

Index Terms

Gaussian belief propagation, two-dimensional channels, multi-access interference, iterative processing.

I. Introduction

Coding and equalization for two-dimensional intersymbol interference (2D ISI) channels has attracted considerable interest during the last decade, due to recent developments in the area of optical and magnetic storage systems. Since the materials used in today's recording technology are expected to reach their physical limits in the near future, new systems using multi-track optical storage [1] and holographic storage [2] have been proposed. Multiple-input multiple-output (MIMO) communication systems over frequency selective fading channels can also be viewed as an instance of 2D ISI [3]. In this paper, we will consider the uplink of a cellular network, where 2D ISI occurs as a result of intercell interference. This problem was first treated in [4], where the presence of adjacent cells is modeled as a multiple-access interference term and Shannon-theoretic limits are obtained.

Maximum-likelihood detection is generally unfeasible on 2D ISI channels, because this would require to use a Viterbi algorithm [5] where each row of data is treated as an input symbol, leading to a huge input alphabet [6]. Therefore several suboptimal detection approaches have been proposed in the literature.

A first family of detection algorithms uses 2D filtering techniques to equalize the 2D ISI channel. The minimum mean-square error (MMSE) equalizer and the decision feedback equalizer (DFE), which are well-known techniques used to equalize 1D channels with memory, have been adapted for 2D ISI channels in [7] and [8], respectively. The obtained results are satisfactory as long as the ISI is not too severe.

A second family of 2D equalizers process independently each row of received observations by taking into account symbols estimated in neighboring rows. The decision feedback Viterbi algorithm (DF-VA) runs a Viterbi algorithm for each row using hard decision feedback from previous rows [9]. The iterative multistrip (IMS) algorithm uses the same idea, but in a more elaborate manner in order to achieve better performances [10]-[11]. The current row of observations is associated with a finite number of previous rows to form a strip. Each row of observations is then processed independently with a soft-input soft-output (SISO) BCJR equalizer [12], generating soft decisions

The author is with INSTITUT TELECOM, TELECOM SudParis, departement CITI, UMR-CNRS 5157, 91011 Evry Cedex, France (e-mail: frederic.lehmann@it-sudparis.eu).

Phone: (+33) 1 60 76 46 33. Fax: (+33) 1 60 76 44 33

Manuscript received September 7, 2011, revised January 16, 2012 and March 12, 2012.

for all the rows of symbols in the corresponding strip at a time. Since the strips corresponding to neighboring rows overlap, multiple soft decisions are generated for each symbol. A factor graph representation of the equality constraints between the symbols in overlapping rows can be constructed and used to derive an iterative message passing algorithm. Thus each iteration runs the BCJR equalizer across the strips with soft decision feedback from the previous iteration. A similar iterative equalizer has also appeared in [6], where the factor graph approach based on equality constraints between strips is replaced by a trellis representation of vector symbols along the columns. The stand-alone performances of the IMS equalizer is generally within 2 dB of the optimal performances. Even better results were obtained in [13] by alternating row and column equalization and by processing more than one row (or column) of observations per strip. However, the performance improvements came at the price of a higher computational complexity per symbol.

Recently, a third class of algorithms relying on belief propagation (BP) have appeared in the literature. We assume that the reader is familiar with the factor graph framework [14] . In [15]-[17], the *a posteriori* probability mass function (pmf) of the information symbols is factorized in a straighforward manner, assuming that the ambient noise is white and the symbols are independent. The resulting bipartite factor graph relates variable nodes corresponding to the symbols with function nodes corresponding to the equations relating the noisy observations to the symbols. Standard belief propagation is then applied to infer the unobserved symbols [18]. However, since such a factor graph typically contains many short cycles, belief propagation may fail to converge. To circumvent this problem, a different factorization of the *a posteriori* pmf of the information symbols, combined with generalized belief propagation (GBP) [19] was proposed in [20].

In this paper, we consider the application of Gaussian belief propagation (GaBP) [21]-[26] to the second family of 2D equalizers. For the sake of simplicity, we will restrict ourselves to processing the rows of observations independently, as in the IMS equalizer [10]. A straightforward generalization would allow the processing of multiple rows of observations per strip as in [13]. In order to apply GaBP, me must assume that all the messages exchanged on the factor graph representation of the 2D equalization problem are Gaussian distributions. This is obviously an approximation since this corresponds to replacing the natural discrete state-space of unobserved information symbols by a continuous state-space. However, the performance loss of iterative equalization using the GaBP-based equalizer instead of the IMS equalizer is small. The proposed approach has a number of advantages. The complexity of the GaBP-based equalizer does not depend on the size of the constellation used for modulation. Moreover, as we shall see, the complexity of the GaBP-based equalizer is not exponential in the size of the state-space as for the IMS equalizer. Therefore, when the spatial interference is not restricted to the immediately adjacent cells, a substantial reduction in computational complexity is obtained.

The main technical contributions of this work are

- A state-space model suitable for the uplink of hexagonal cellular networks with full frequency reuse, where intercell interference is regarded as causal 2D ISI affecting each base station.
- A 2D interference cancellation method based on a Gaussian approximation of all the messages used in belief propagation on a factor graph representation of the proposed state-space model.
- The use of higher order modulations and the cancellation of interference, not limited to the nearest neighboring cells, becomes feasible with an acceptable loss of power efficiency, thanks to a lower computational complexity of the proposed scheme relatively to regular BP and BCJR-like algorithms.

We have presented a preliminary version of the GaBP-based equalizer, suitable for binary phase shift keying (BPSK), in the conference paper [27]. In this paper, we provide a unified presentation of the proposed equalization technique along with a generalization to higher order modulations. We also mention that the resulting 2D equalizer bears some resemblance to the algorithm by Guo and Ping, suitable only for 1D ISI channels [28].

Throughout the paper, bold letters indicate vectors and matrices, while $\mathbf{I}_m$ and $\mathbf{0}_{m\times n}$ denote the $m \times m$ identity matrix and the $m \times n$ all-zero matrix, respectively. $\mathcal{N}(\mathbf{x} : \mathbf{m}, \mathbf{P})$ denotes a complex Gaussian distribution of the variable $\mathbf{x}$, with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$.

This paper is organized as follows. First, Sec. II models intercell interference as a 2D ISI problem, which is in turn cast in a state-space form. In Sec. III, we introduce our GaBP-based SISO equalizer processing a single row of observations. Sec. IV is devoted to iterative equalization for 2D ISI channels. Finally, in Sec. V, the performances of the proposed algorithm are assessed through numerical simulations and compared with existing message passing methods.

II. System Model

A. Network Model

We consider a cellular network of $I \times J$ cells. Within each cell, the users are assumed orthogonal (either in time, frequency or code). Therefore, without loss of generality, we can consider a single user per cell. In the cell indexed by $(i,j), 0 \leq i \leq I-1, 0 \leq j \leq J-1$, the user transmits an i.i.d. (independently and identically distributed) data symbol $b_{i,j}$, drawn from the discrete alphabet (constellation) $\mathcal{A} = \{a_0, \ldots, a_{|\mathcal{A}|-1}\}$, to the intended base station (BS). With the standard assumption that the multi-access interference is confined to adjacent cells, the received observation for cell (i,j) can be written as

$$y_{i,j} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} h_{i,j}^{m,n} b_{i-m,j-n} + n_{i,j} \tag{1}$$

where $\{h_{i,j}^{m,n}, 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$ are the coefficients of the 2D multi-access interference channel and $n_{i,j}$ is a complex additive white Gaussian noise (AWGN) term with zero mean and variance σ^2. In Appendix A, we show how the 2D channel, which is a priori non-causal, can be converted to the causal model adopted in (1).

We adapt the network model proposed by Aktas et al. [16], which is able to take into account the effect of path loss and fading, to the standard hexagonal network structure depicted in Fig. 1. The 2D channel coefficients $\{h_{i,j}^{m,n}\}$ are independently distributed according to a zero-mean complex Gaussian distribution.

The first channel model, denoted by channel A, is representative of an hexagonal network with interference truncated to the 6 nearest neighboring cells. Therefore, the matrix of 2D channel coefficients for cell (i,j) is such that $M = 3$, $N = 3$, and has the form

$$\text{Channel A: } [h_{i,j}^{m,n}] = \begin{bmatrix} h_{i,j}^{0,0} & h_{i,j}^{0,1} & 0 \\ h_{i,j}^{1,0} & h_{i,j}^{1,1} & h_{i,j}^{1,2} \\ 0 & h_{i,j}^{2,1} & h_{i,j}^{2,2} \end{bmatrix}$$

Moreover, as per the model by Aktas et al., the variances of the 2D channel coefficients are independent of the cell index (i,j), and are given in matrix form by

$$\text{Channel A: } \left[E\left(|h_{i,j}^{m,n}|^2\right)\right] = \begin{bmatrix} a & a & 0 \\ a & 1 & a \\ 0 & a & a \end{bmatrix}, \quad \forall (i,j)$$

where a is parameter of choice.

The second channel model, denoted by channel B, is representative of an hexagonal network with interference truncated to the 18 nearest neighboring cells. Therefore, the matrix of 2D channel coefficients for cell (i,j) is such that $M = 5$, $N = 5$, and has the form

$$\text{Channel B: } [h_{i,j}^{m,n}] = \begin{bmatrix} h_{i,j}^{0,0} & h_{i,j}^{0,1} & h_{i,j}^{0,2} & 0 & 0 \\ h_{i,j}^{1,0} & h_{i,j}^{1,1} & h_{i,j}^{1,2} & h_{i,j}^{1,3} & 0 \\ h_{i,j}^{2,0} & h_{i,j}^{2,1} & h_{i,j}^{2,2} & h_{i,j}^{2,3} & h_{i,j}^{2,4} \\ 0 & h_{i,j}^{3,1} & h_{i,j}^{3,2} & h_{i,j}^{3,3} & h_{i,j}^{3,4} \\ 0 & 0 & h_{i,j}^{4,2} & h_{i,j}^{4,3} & h_{i,j}^{4,4} \end{bmatrix}$$

Again, as per the model by Aktas et al., the variances of the 2D channel coefficients are independent of the cell index (i,j), and are given in matrix form by

$$\text{Channel B: } \left[E\left(|h_{i,j}^{m,n}|^2\right)\right] = \begin{bmatrix} b & b & b & 0 & 0 \\ b & a & a & b & 0 \\ b & a & 1 & a & b \\ 0 & b & a & a & b \\ 0 & 0 & b & b & b \end{bmatrix}, \quad \forall (i,j)$$

where a and b are parameters of choice.

More general channels with longer memory are easily obtained by taking into account interference from neighboring cells even further away from each BS.

B. State-space Model

For the i-th line of observations $\{y_{i,k}, 0 \leq k \leq J-1\}$, we define the state $\mathbf{x}_k$ of size $MN \times 1$

$$\mathbf{x}_k = \begin{bmatrix} b_{i-(M-1),k-(N-1)} \\ \vdots \\ b_{i,k-(N-1)} \\ \vdots \\ b_{i-(M-1),k} \\ \vdots \\ b_{i,k} \end{bmatrix}, \tag{2}$$

which consists in raster scanning columnwise the data symbols illustrated in Fig. 2. Let us define the vector of 2D ISI coefficients of size $MN \times 1$ for cell (i,j)

$$\mathbf{h}_{i,j} = \begin{bmatrix} h_{i,j}^{M-1,N-1} \\ \vdots \\ h_{i,j}^{0,N-1} \\ \vdots \\ h_{i,j}^{M-1,0} \\ \vdots \\ h_{i,j}^{0,0} \end{bmatrix}, \tag{3}$$

then we obtain the following state-space representation for the i-th line of observations

$$\begin{cases} \mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\mathbf{i}_k \\ y_{i,k} = \mathbf{h}_{i,k}^T \mathbf{x}_k + n_{i,k}, \quad 0 \leq k \leq J-1, \end{cases} \tag{4}$$

where the new input vector at instant k is defined as

$$\mathbf{i}_k = \begin{bmatrix} b_{i-(M-1),k} \\ \vdots \\ b_{i,k} \end{bmatrix}.$$

The state transition matrices are given by

$$\mathbf{F} = \begin{bmatrix} \mathbf{0}_{N-1\times 1} & \mathbf{I}_{N-1} \\ 0 & \mathbf{0}_{1\times N-1} \end{bmatrix} \otimes \mathbf{I}_M$$

and

$$\mathbf{G} = \begin{bmatrix} \mathbf{0}_{N-1\times 1} \\ 1 \end{bmatrix} \otimes \mathbf{I}_M,$$

where the symbol $\otimes$ denotes the Kronecker product.

The *a posteriori* pmf of the input vectors and the 2D ISI states corresponding to the i-th line can be factorized as follows (see Appendix B)

$$p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J} | y_{i,1:J}) \propto p(\mathbf{x}_0) \prod_{k=1}^{J} p(\mathbf{i}_k) p(\mathbf{x}_k | \mathbf{i}_k, \mathbf{x}_{k-1}) p(y_{i,k} | \mathbf{x}_k). \tag{5}$$

The corresponding factor graph is depicted in Fig. 3 [14], where the function node f_k represents the factor $p(\mathbf{x}_k|\mathbf{i}_k, \mathbf{x}_{k-1})$ and the function node g_k represents the factor $p(y_{i,k}|\mathbf{x}_k)$. From (4) we have

$$\begin{cases} f_k = p(\mathbf{x}_k|\mathbf{i}_k, \mathbf{x}_{k-1}) = \delta\left(\mathbf{x}_k - (\mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\mathbf{i}_k)\right) \\ g_k = p(y_{i,k}|\mathbf{x}_k) = \mathcal{N}(y_{i,k} : \mathbf{h}_{i,k}^T \mathbf{x}_k, \sigma^2). \end{cases}$$

III. GABP-BASED EQUALIZATION

In this section, we derive a low-complexity 2D equalizer based on GaBP [21]-[26] by approximating all the messages, exchanged by the sum-product algorithm [14] on the factor graph depicted in Fig. 3, as Gaussian distributions.

A. Prior distributions

First, we seek a Gaussian approximation of the prior distributions $p(\mathbf{x}_0)$ and $p(\mathbf{i}_k)$, $k = 1, \ldots, J$ appearing in the factor graph.

We assume that $p(\mathbf{x}_0)$ is known and can be approximated as the Gaussian distribution $\mathcal{N}(\mathbf{x}_0 : \hat{\mathbf{x}}_0, \mathbf{P}_0)$, whose parameters are chosen as $\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$ and $\mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^H]$.

Assuming that symbol-wise *a priori* log-likelihood ratios (LLRs) are available for the components of $\mathbf{i}_k$, i.e.

$$\begin{cases} L_a^{a_q}(b_{i-(M-1),k}) = \ln \dfrac{p(b_{i-(M-1),k} = a_q)}{p(b_{i-(M-1),k} = a_0)}, \quad q = 0, \ldots, |\mathcal{A}| - 1 \\ \qquad\qquad \vdots \\ L_a^{a_q}(b_{i,k}) = \ln \dfrac{p(b_{i,k} = a_q)}{p(b_{i,k} = a_0)}, \quad q = 0, \ldots, |\mathcal{A}| - 1 \end{cases} \tag{6}$$

a Gaussian approximation of $p(\mathbf{i}_k)$ of the form $\mathcal{N}(\mathbf{i}_k : \hat{\mathbf{i}}_k, \mathbf{Q}_k)$ is sought. Again, the standard solution consists in matching the parameters $\hat{\mathbf{i}}_k$ and $\mathbf{Q}_k$ to the mean and the covariance matrix of the discrete-valued input vector $\mathbf{i}_k$, respectively. Using the fact that the data symbols are assumed independent yields [29]

$$\hat{\mathbf{i}}_k = E[\mathbf{i}_k] = \begin{bmatrix} m_a(b_{i-(M-1),k}) \\ \vdots \\ m_a(b_{i,k}) \end{bmatrix}$$

where

$$m_a(b_{i-m,k}) = E\left[b_{i-m,k}\right], \quad m = 0, \ldots, M-1,$$

and

$$\begin{aligned} \mathbf{Q}_k &= E[(\mathbf{i}_k - \hat{\mathbf{i}}_k)(\mathbf{i}_k - \hat{\mathbf{i}}_k)^H] \\ &= \operatorname{diag}\left(\sigma_a(b_{i-(M-1),k})^2, \ldots, \sigma_a(b_{i,k})^2\right), \end{aligned} \tag{7}$$

where

$$\sigma_a(b_{i-m,k})^2 = E\left[|b_{i-m,k}|^2\right] - |m_a(b_{i-m,k})|^2, \quad m = 0, \ldots, M-1.$$

Using (6), we can easily show that for $m = 0, \ldots, M-1$

$$\begin{aligned} E\left[b_{i-m,k}\right] &= \sum_{q=0}^{|\mathcal{A}|-1} a_q \frac{e^{L_a^{a_q}(b_{i-m,k})}}{\sum_{p=0}^{|\mathcal{A}|-1} e^{L_a^{a_p}(b_{i-m,k})}} \\ E\left[|b_{i-m,k}|^2\right] &= \sum_{q=0}^{|\mathcal{A}|-1} |a_q|^2 \frac{e^{L_a^{a_q}(b_{i-m,k})}}{\sum_{p=0}^{|\mathcal{A}|-1} e^{L_a^{a_p}(b_{i-m,k})}}. \end{aligned}$$

Remark 3.1: For BPSK modulation, let $a_0 = +1$ and $a_1 = -1$, we obtain the simplified expressions in [27]

$$\begin{aligned} m_a(b_{i-m,k}) &= -\tanh\left(\frac{L_a^{a_1}(b_{i-m,k})}{2}\right), \\ \sigma_a(b_{i-m,k})^2 &= 1 - m_a(b_{i-m,k})^2, \quad m = 0, \ldots, M-1 \end{aligned}$$

Remark 3.2: In (7), the diagonal elements of $\mathbf{Q}_k$ tend to 0 when the *a priori* knowledge about the data symbols becomes very reliable. In order to ensure numerical stability of GaBP, the diagonal elements of $\mathbf{Q}_k$ are constraint to be greater than a threshold, say $\geq 10^{-3}$.

B. Forward pass

Let $\mu_{u\to v}(.)$ be the message sent by node u to node v in the factor graph. Fig. 4 illustrates the notation on a portion of the factor graph around the 2D ISI state variable $\mathbf{x}_k$.

Assuming that $\mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1})$ is parameterized by a Gaussian distribution such that

$$\mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_{k-1} : \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}), \tag{8}$$

the forward pass consists in computing a recursive expression for $\mu_{\mathbf{x}_k\to f_{k+1}}(\mathbf{x}_k)$.

Let us first apply the sum-product rule at the function node f_k, we obtain

$$\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k) = \int\int p(\mathbf{x}_k|\mathbf{i}_k, \mathbf{x}_{k-1}) p(\mathbf{i}_k) \mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1}) d\mathbf{i}_k d\mathbf{x}_{k-1} \tag{9}$$

We show in Appendix C that $\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k)$ admits the following closed form expression

$$\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k) \propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}), \tag{10}$$

where

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{G}\hat{\mathbf{i}}_k \\ \mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^H + \mathbf{G}\mathbf{Q}_k\mathbf{G}^H. \end{cases}$$

Now, applying the sum-product rule to the variable node $\mathbf{x}_k$ yields

$$\begin{aligned} \mu_{\mathbf{x}_k\to f_{k+1}}(\mathbf{x}_k) &= \mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k)\mu_{g_k\to\mathbf{x}_k}(\mathbf{x}_k) \\ &\propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\mathcal{N}(y_{i,k} : \mathbf{h}_{i,k}^T\mathbf{x}_k, \sigma^2). \end{aligned}$$

We recognize the correction step of the well-known Kalman filter [30], therefore

$$\mu_{\mathbf{x}_k\to f_{k+1}}(\mathbf{x}_k) \propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}), \tag{11}$$

where

$$\begin{cases} \mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{h}_{i,k}^*(\mathbf{h}_{i,k}^T\mathbf{P}_{k|k-1}\mathbf{h}_{i,k}^* + \sigma^2)^{-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(y_{i,k} - \mathbf{h}_{i,k}^T\hat{\mathbf{x}}_{k|k-1}) \\ \mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{h}_{i,k}^T\mathbf{P}_{k|k-1}. \end{cases}$$

C. Backward pass

We are now looking for recursions analog to those found in Sec. III-B for the messages in the factor graph propagated in the backward direction. Let us apply the sum-product rule to the function node f_{k+1},

$$\begin{aligned} \mu_{f_{k+1}\to\mathbf{x}_k}(\mathbf{x}_k) &= \int\int p(\mathbf{x}_{k+1}|\mathbf{i}_{k+1}, \mathbf{x}_k) p(\mathbf{i}_{k+1}) \mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1}) d\mathbf{i}_{k+1} d\mathbf{x}_{k+1} \\ &= \int \left(\int p(\mathbf{x}_{k+1}|\mathbf{i}_{k+1}, \mathbf{x}_k) p(\mathbf{i}_{k+1}) d\mathbf{i}_{k+1} \right) \mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1}) d\mathbf{x}_{k+1} \\ &= \int p(\mathbf{x}_{k+1}|\mathbf{x}_k) \mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1}) d\mathbf{x}_{k+1}. \end{aligned} \tag{12}$$

We easily recognize that (12) corresponds to the backwards prediction step in a two-filter Kalman smoother [31], therefore, $\mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1}) \propto p(y_{i,k+1:J}|\mathbf{x}_{k+1})$. It follows that $\mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1})$ is a likelihood, which in general cannot be assimilated to a Gaussian probability density of $\mathbf{x}_{k+1}$, as required for GaBP. However, from Bayes rule we have $p(\mathbf{x}_{k+1}|y_{i,k+1:J}) \propto p(\mathbf{x}_{k+1})p(y_{i,k+1:J}|\mathbf{x}_{k+1})$. Thus $p(\mathbf{x}_{k+1})\mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1})$ can be parameterized by a Gaussian density of $\mathbf{x}_{k+1}$ such that

$$p(\mathbf{x}_{k+1})\mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1}) = \mathcal{N}(\mathbf{x}_{k+1} : \hat{\mathbf{x}}_{k+1|k+1:J}, \mathbf{P}_{k+1|k+1:J}). \tag{13}$$

We must rearrange (12) so that (13) appears explicitly in the recursion.

First, we express the forward dynamics $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ as a function of the backward dynamics $p(\mathbf{x}_k|\mathbf{x}_{k+1})$ using Bayes rule

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \frac{p(\mathbf{x}_k|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1})}{p(\mathbf{x}_k)}. \tag{14}$$

We can show that the prior distribution of the 2D ISI states and the corresponding backward dynamics are Gaussian of the form

$$\begin{cases} p(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_k, \mathbf{\Pi}_k) \\ p(\mathbf{x}_k|\mathbf{x}_{k+1}) = \mathcal{N}(\mathbf{x}_k : \tilde{\mathbf{F}}_{k+1}\mathbf{x}_{k+1} + \mathbf{c}_{k+1}, \tilde{\mathbf{Q}}_{k+1}) \end{cases} \tag{15}$$

(see Appendix D for the demonstration and the expression of the parameters $\hat{\mathbf{x}}_k, \mathbf{\Pi}_k, \tilde{\mathbf{F}}_{k+1}, \mathbf{c}_{k+1}, \tilde{\mathbf{Q}}_{k+1}$).

Then, injecting (14) into (12) leads to

$$\begin{aligned} p(\mathbf{x}_k)\mu_{f_{k+1}\to\mathbf{x}_k}(\mathbf{x}_k) &= \int p(\mathbf{x}_k|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1})\mu_{\mathbf{x}_{k+1}\to f_{k+1}}(\mathbf{x}_{k+1})d\mathbf{x}_{k+1} \\ &= \int \mathcal{N}(\mathbf{x}_k : \tilde{\mathbf{F}}_{k+1}\mathbf{x}_{k+1} + \mathbf{c}_{k+1}, \tilde{\mathbf{Q}}_{k+1})\mathcal{N}(\mathbf{x}_{k+1} : \hat{\mathbf{x}}_{k+1|k+1:J}, \mathbf{P}_{k+1|k+1:J})d\mathbf{x}_{k+1}. \end{aligned} \tag{16}$$

Noting the the last equality in (16) is the integral appearing in the (backward) prediction step of Kalman filtering [30], we readily have

$$p(\mathbf{x}_k)\mu_{f_{k+1}\to\mathbf{x}_k}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k+1:J}, \mathbf{P}_{k|k+1:J}), \tag{17}$$

where

$$\begin{cases} \hat{\mathbf{x}}_{k|k+1:J} = \tilde{\mathbf{F}}_{k+1}\hat{\mathbf{x}}_{k+1|k+1:J} + \mathbf{c}_{k+1} \\ \mathbf{P}_{k|k+1:J} = \tilde{\mathbf{F}}_{k+1}\mathbf{P}_{k+1|k+1:J}\tilde{\mathbf{F}}_{k+1}^H + \tilde{\mathbf{Q}}_{k+1}. \end{cases}$$

To complete the backward pass, we apply the sum-product rule to the variable node $\mathbf{x}_k$

$$\begin{aligned} p(\mathbf{x}_k)\mu_{\mathbf{x}_k\to f_k}(\mathbf{x}_k) &= p(\mathbf{x}_k)\mu_{f_{k+1}\to\mathbf{x}_k}(\mathbf{x}_k)\mu_{g_k\to\mathbf{x}_k}(\mathbf{x}_k) \\ &\propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k+1:J}, \mathbf{P}_{k|k+1:J})\mathcal{N}(y_{i,k} : \mathbf{h}_{i,k}^T\mathbf{x}_k, \sigma^2). \end{aligned}$$

Again, we recognize the (backward) correction step of the well-known Kalman filter [30], therefore

$$p(\mathbf{x}_k)\mu_{\mathbf{x}_k\to f_k}(\mathbf{x}_k) \propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k:J}, \mathbf{P}_{k|k:J}), \tag{18}$$

where

$$\begin{cases} \tilde{\mathbf{K}}_k = \mathbf{P}_{k|k+1:J}\mathbf{h}_{i,k}^*(\mathbf{h}_{i,k}^T\mathbf{P}_{k|k+1:J}\mathbf{h}_{i,k}^* + \sigma^2)^{-1} \\ \hat{\mathbf{x}}_{k|k:J} = \hat{\mathbf{x}}_{k|k+1:J} + \tilde{\mathbf{K}}_k(y_{i,k} - \mathbf{h}_{i,k}^T\hat{\mathbf{x}}_{k|k+1:J}) \\ \mathbf{P}_{k|k:J} = \mathbf{P}_{k|k+1:J} - \tilde{\mathbf{K}}_k\mathbf{h}_{i,k}^T\mathbf{P}_{k|k+1:J}. \end{cases}$$

D. Smoothing pass

The sum-product algorithm computes the *a posteriori* marginal probability distribution of the 2D ISI state $\mathbf{x}_k$ as the product of all incoming messages to variable node $\mathbf{x}_k$ in Fig. 4

$$p(\mathbf{x}_k|y_{i,1:J}) \propto \frac{\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k) \times p(\mathbf{x}_k)\mu_{f_{k+1}\to\mathbf{x}_k}(\mathbf{x}_k)\mu_{g_k\to\mathbf{x}_k}(\mathbf{x}_k)}{p(\mathbf{x}_k)}$$

which, according to (10), (18) and (15), can be rewritten as

$$p(\mathbf{x}_k|y_{i,1:J}) \propto \frac{\mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|k:J}, \mathbf{P}_{k|k:J})}{\mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_k, \mathbf{\Pi}_k)}.$$

After straightforward algebraic manipulations, we obtain the following simplification

$$p(\mathbf{x}_k|y_{i,1:J}) \propto \mathcal{N}(\mathbf{x}_k : \hat{\mathbf{x}}_{k|1:J}, \mathbf{P}_{k|1:J}), \tag{19}$$

where

$$\begin{cases} \mathbf{P}_{k|1:J} = \left[\mathbf{P}_{k|k-1}^{-1} + \mathbf{P}_{k|k:J}^{-1} - \mathbf{\Pi}_k^{-1}\right]^{-1} \\ \hat{\mathbf{x}}_{k|1:J} = \mathbf{P}_{k|1:J}^{-1}\left[\mathbf{P}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1} + \mathbf{P}_{k|k:J}^{-1}\hat{\mathbf{x}}_{k|k:J} - \mathbf{\Pi}_k^{-1}\hat{\mathbf{x}}_k\right]. \end{cases} \tag{20}$$

E. Computation of symbol-wise extrinsic log-likelihood ratios

From (2), the M last coordinates of $\mathbf{x}_k$ correspond to the data vector $\mathbf{i}_k = [b_{i-(M-1),k}, \ldots, b_{i,k}]^T$. Therefore, the *a posteriori* probability distribution of $b_{i-m,k}$, seen as a complex random variable, has the form

$$p(b_{i-m,k}|y_{i,1:J}) \propto \mathcal{N}(b_{i-m,k} : m_p(b_{i-m,k}), \sigma_p(b_{i-m,k})^2), \quad m = 0, \ldots, M-1, \tag{21}$$

where the mean $m_p(b_{i-m,k})$ and the variance $\sigma_p(b_{i-m,k})^2$, $m = 0, \ldots, M-1$ are easily extracted from (20) by marginalization. Converting $b_{i-m,k}$ back to a discrete random variable in $\mathcal{A}$, the expression of the symbol-wise *a posteriori* log-likelihood ratio is obtained for $m = 0, \ldots, M-1$ and $a_q \in \mathcal{A}$ as [33]

$$\begin{aligned} L_p^{a_q}(b_{i-m,k}|y_{i,1:J}) &= \ln \frac{p(b_{i-m,k} = a_q|y_{i,1:J})}{p(b_{i-m,k} = a_0|y_{i,1:J})} \\ &= \ln \frac{\frac{1}{\pi\sigma_p(b_{i-m,k})^2} e^{-\frac{|a_q - m_p(b_{i-m,k})|^2}{\sigma_p(b_{i-m,k})^2}}}{\frac{1}{\pi\sigma_p(b_{i-m,k})^2} e^{-\frac{|a_0 - m_p(b_{i-m,k})|^2}{\sigma_p(b_{i-m,k})^2}}}. \end{aligned}$$

However, in message passing algorithms, the messages of interest are symbol-wise *extrinsic* log-likelihood ratios obtained for $m = 0, \ldots, M-1$ and $a_q \in \mathcal{A}$ as

$$L_e^{a_q}(b_{i-m,k}|y_{i,1:J}) = w \left(\ln \frac{\frac{1}{\pi\sigma_p(b_{i-m,k})^2} e^{-\frac{|a_q - m_p(b_{i-m,k})|^2}{\sigma_p(b_{i-m,k})^2}}}{\frac{1}{\pi\sigma_p(b_{i-m,k})^2} e^{-\frac{|a_0 - m_p(b_{i-m,k})|^2}{\sigma_p(b_{i-m,k})^2}}} - \ln \frac{\frac{1}{\pi\sigma_a(b_{i-m,k})^2} e^{-\frac{|a_q - m_a(b_{i-m,k})|^2}{\sigma_a(b_{i-m,k})^2}}}{\frac{1}{\pi\sigma_a(b_{i-m,k})^2} e^{-\frac{|a_0 - m_a(b_{i-m,k})|^2}{\sigma_a(b_{i-m,k})^2}}} \right), \tag{22}$$

where the scaling factor $0 \leq w \leq 1$. The reason for introducing the scaling factor is that suboptimum detectors tend to overestimate the reliability of their soft-outputs, which adversely affects the convergence of iterative methods using such detectors as building blocks [34]-[35]. We apply a well-known solution to this problem, which consists in extrinsic information weighting [34]-[35].

Remark 3.3: For BPSK modulation, let $a_0 = +1$ and $a_1 = -1$, we obtain the simplified expressions

$$L_p^{a_1}(b_{i-m,k}|y_{i,1:J}) = -\frac{4\,\mathrm{Re}(m_p(b_{i-m,k}))}{\sigma_p(b_{i-m,k})^2}, \quad m = 0, \ldots, M-1$$

and

$$L_e^{a_1}(b_{i-m,k}|y_{i,1:J}) = -4w \left(\frac{\mathrm{Re}(m_p(b_{i-m,k}))}{\sigma_p(b_{i-m,k})^2} - \frac{\mathrm{Re}(m_a(b_{i-m,k}))}{\sigma_a(b_{i-m,k})^2} \right), \quad m = 0, \ldots, M-1.$$

IV. GaBP-based iterative equalization

A. Proposed iterative method

The algorithm described in the previous section is suitable only for processing one line of observations. Moreover, it can be seen from (22) that the GaBP equalizer corresponding to the i-th line of observations generates soft outputs for the current line, but also for the $M-1$ previous lines of data symbols. Therefore, once all the lines of observations have been processed, M soft outputs are available for each data symbol.

In order to describe the equalization process in the context of a 2D cellular network, we need to construct the complete factor graph corresponding to all the lines of observations. The complete factor graph is represented in Fig. 5, for channel A (hence $M = 3$). The variable nodes corresponding to the data symbols $\{b_{i,j}\}$ and the noisy observations $\{y_{i,j}\}$ of all cells have been linewise raster scanned. In Fig. 5, when processing the i-th line of observations at any time instant k, the corresponding input vector

$$\mathbf{i}_k = \begin{bmatrix} b_{i-(M-1),k} \\ \vdots \\ b_{i,k} \end{bmatrix}$$

receives input *a priori* log-likelihood ratios A_0 for $b_{i,k}$, A_1 for $b_{i-1,k}$, ..., A_{M-1} for $b_{i-(M-1),k}$, where A_m is a shorthand notation for the quantities $L_a^{a_q}(b_{i-m,k})$ defined in (6). Then, output *a posteriori* log-likelihood ratios E_0

1) Initialize the iteration index: $l = 0$
2) Initialize the messages sent by the GaBP equalization subgraphs towards the data symbol nodes: $E_0(l) = 0$, $E_1(l) = 0$, ..., $E_{M-1}(l) = 0$
3) for $l = 1 : N_t$
   - The messages $A_0(l)$, $A_1(l)$, ..., $A_{M-1}(l)$, sent by a data symbol node towards a GaBP equalization subgraph are the sum of all incoming messages, except the message coming from that GaBP equalization subgraph.
   - The messages $E_0(l)$, $E_1(l)$, ..., $E_{M-1}(l)$, sent by a GaBP equalization subgraph towards a data symbol node, are given by (22).

TABLE I
GABP-BASED ITERATIVE EQUALIZATION

for $b_{i,k}$, E_1 for $b_{i-1,k}$, ..., E_{M-1} for $b_{i-M-1,k}$ are produced, where E_m is a shorthand notation for the quantities $L_e^{a_q}(b_{i-m,k}|y_{i,1:J})$ defined in (22).

The algorithm of Table I summarizes the proposed iterative procedure. Since the complete factor graph is loopy, we perform N_t iterations, where one iteration corresponds to one round of message computation for all data symbol nodes, followed by GaBP equalization applied to all the lines of observations (see Sec. III). Hence the name GaBP-based iterative equalization for the proposed algorithm.

Since we chose to process one line of observations at a time, the receiver will have the same structure as the IMS decoder originally introduced in [10], except that the BCJR equalizers are replaced with GaBP equalizers. From a practical point of view, GaBP-based iterative equalization is a form of distributed message passing between the base stations. To be more specific, for each iteration, during the processing of one line of observations, each base station performs local processing of the measurement from the mobile device located in its cell. Each base station then communicates soft information to $M - 1$ neighboring base stations. From an implementation point of view, high bandwidth connections, such fast fiber links, are required for sharing information between neighboring base stations, as explained in [38].

B. Convergence analysis

We use extrinsic information transfer (EXIT) charts [36] to analyze the convergence properties of the proposed iterative equalizer. For ease of exposition, we will restrict ourselves to BPSK modulation. A generalization to higher order modulations would be obtained by adapting the technique proposed in [37].

Ignoring the presence of cycle in the graph, we can assume that all quantities A_m and E_m, $0 \leq m \leq M-1$, at any time instant are i.i.d. (identically and idependently distributed). Following [36], we assume that A_m is Gaussian distributed with variance $\sigma_{A_m}^2$ and mean $\sigma_{A_m}^2/2$, $0 \leq m \leq M-1$. Then, the input mutual information between the *a priori* log-likelihood ratios A_m and the corresponding binary data symbols, for $0 \leq m \leq M-1$, has the form [36]

$$\begin{cases} IA_0 = J(\sigma_{A_0}) \\ IA_1 = J(\sigma_{A_1}) \\ \vdots \\ IA_{M-1} = J(\sigma_{A_{M-1}}) \end{cases}$$

where

$$J(\sigma) = 1 - \int_{-\infty}^{+\infty} \frac{e^{-(\xi-\sigma^2/2)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \log_2(1+e^{-\xi}) d\xi.$$

For a fixed value of signal-to-noise ratio (SNR), we collect the *a posteriori* log-likelihood ratios after one round of iterative GaBP equalization, obtained by performing Monte-Carlo simulations on a cellular network with $I =$

$J = 20$, with random 2D channel coefficients drawn according to the channel model in Sec. II-A. Moreover, *a priori* log-likelihood ratios, A_m, are drawn from a real Gaussian distribution with mean $\sigma_{A_m}^2/2$ and variance $\sigma_{A_m}^2$, $0 \leq m \leq M-1$. Using the histogram method proposed in [36], the output mutual information between the *a posteriori* log-likelihood ratios E_m and the corresponding binary data symbols, for $0 \leq m \leq M-1$, can be computed. The results can be presented in the form of mapping functions $T_0(.)$, $T_1(.)$, ..., $T_{M-1}(.)$ as follows

$$\begin{cases} IE_0 = T_0(IA_1, IA_2, \ldots, IA_{M-1}, SNR) \\ IE_1 = T_1(IA_1, IA_2, \ldots, IA_{M-1}, SNR) \\ \quad \vdots \\ IE_{M-1} = T_{M-1}(IA_1, IA_2, \ldots, IA_{M-1}, SNR). \end{cases}$$

The evolution of the input/output mutual information and the predicted bit-error rate (BER), P_b, for the algorithm of Table I is summarized in Table II [36].

1) Initialize the iteration index: $l = 0$
2) $IE_0(l) = 0$, $IE_1(l) = 0$, ..., $IE_{M-1}(l) = 0$
3) for $l = 1 : N_t$

$$IA_0(l) = J\left(\sqrt{\sum_{m\neq 0} J^{-1}(IE_m(l-1))^2}\right)$$

$$IA_1(l) = J\left(\sqrt{\sum_{m\neq 1} J^{-1}(IE_m(l-1))^2}\right)$$

$$\vdots$$

$$IA_{M-1}(l) = J\left(\sqrt{\sum_{m\neq M-1} J^{-1}(IE_m(l-1))^2}\right)$$

$$IE_0(l) = T_0\left(IA_1(l), IA_2(l), \ldots, IA_{M-1}(l), SNR\right)$$

$$IE_1(l) = T_1\left(IA_1(l), IA_2(l), \ldots, IA_{M-1}(l), SNR\right)$$

$$\vdots$$

$$IE_{M-1}(l) = T_{M-1}\left(IA_1(l), IA_2(l), \ldots, IA_{M-1}(l), SNR\right)$$

$$P_b(l) = 0.5\,\mathrm{erfc}\left(\frac{\sqrt{\sum_{m=0}^{M-1} J^{-1}(IE_m(l))^2}}{2\sqrt{2}}\right)$$

TABLE II
EVOLUTION OF THE INPUT/OUTPUT MUTUAL INFORMATION AND THE PREDICTED BER

C. Complexity comparison for the proposed and existing methods

For one line of observations, the IMS receiver applies the sum-product algorithm on the factor graph depicted in Fig. 3, assuming that the input vectors $\{\mathbf{i}_k\}$ are discrete-valued. Thus exact Bayesian inference is implemented, because the factor graph has a tree structure [18]. Since the input vectors $\{\mathbf{i}_k\}$ are discrete-valued, the hidden states $\{\mathbf{x}_k\}$ are also discrete-valued and their dynamics can be represented by a trellis with $|\mathcal{A}|^{NM}$ states [5]. The sum-product algorithm then reduces to the Baum-Welch algorithm [39]. A slightly different state-space model for 2D ISI was introduced in [10], where the likelihood of the observations depends on the state transitions instead of the states as per our model. This enables to represent the dynamics of the 2D ISI states with a trellis containing only $|\mathcal{A}|^{(N-1)M}$ states. The sum-product algorithm then becomes equivalent to the BCJR algorithm [12]. Thus, the computational complexity of the IMS receiver is $\mathcal{O}(|\mathcal{A}|^{NM})$ per data symbol and per iteration. Similarly, one

can show that the complexity per data symbol and per iteration of the BP method [15]-[17] is also $\mathcal{O}(|\mathcal{A}|^{NM})$. Therefore these 2D equalization procedures become untractable even for moderate values of M, N and $|\mathcal{A}|$ (say for $M, N > 3$ and $|\mathcal{A}| > 2$).

The proposed iterative method described in Sec. IV-A relies on the GaBP equalizer of Sec. III, which is in essence a Kalman smoother incorporating a priori information. A detailed complexity analysis of the Kalman smoother is provided in [40]. Namely, it is shown that the computational complexity cost per recursion of the Kalman smoother is $\mathcal{O}(p^3)$, where p denotes the size of the state vector. Note that this result holds irrespective of the actual content of the state vector. Remarkably, this implies that the complexity of the proposed method is independent of the size of the constellation used for modulation. Since the size of the state vector given by (2) is MN, the computational complexity of the proposed GaBP-based iterative equalization method is $\mathcal{O}((MN)^3)$ per data symbol and per iteration. Note that unlike the IMS or the BP method, the complexity of the proposed algorithm has only a modest cubically (instead of exponentially) increasing complexity as a function of the size of the state-space.

V. Simulation results

The performance of the proposed schemes is assessed by computer simulations in terms of bit-error rate (BER) versus the normalized signal-to-noise ratio (SNR) defined as

$$SNR = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} E\left[|h_{i,j}^{m,n}|^2\right]}{\sigma^2 \log_2 |\mathcal{A}|},$$

which is independent of the cell index (i, j), according to the network model introduced in Sec. II-A. Unless otherwise stated, iterative equalization is simulated for channel A, with $M = N = 3$, corresponding to an hexagonal network with intercell interference limited to the 6 nearest neighboring cells. Also, we assume that the base stations have perfect knowledge of their 2D channel parameters. The size of the considered cellular network is set to $I = J = 20$. Our experiments showed that the influence of the network size on the bit error rate is not significant. Guard bands, similar to those described in [41], are used to initialize and terminate the equalizer processing each row of observations in a known state. In particular for the GaBP equalizer of Sec. III, the prior distribution of the initial state, $\mathcal{N}(\mathbf{x}_0 : \hat{\mathbf{x}}_0, \mathbf{P}_0)$, can be chosen such that $\hat{\mathbf{x}}_0$ is the known data vector contained in the guard band and $\mathbf{P}_0 = 10^{-3}\mathbf{I}_{MN}$. The weight factor appearing in (22) was set empirically to $w = 0.7$.

A. Convergence analysis using EXIT functions

We study the convergence properties of the proposed GaBP-based iterative equalization technique by applying the EXIT function methodology described in Sec. IV-B. For simplicity, we restrict ourselves to BPSK modulation and channel A (i.e. $M = 3$). Fig. 6-7 (resp. Fig. 8) illustrate the evolution of the input/output mutual information (resp. the predicted BER) as a function of the iteration index, when $a = 0.5$ and $SNR = 10$ dB. We observe that the proposed algorithm converges within $N_t = 5$ iterations. Similar results were obtained for different values of a and SNR.

The EXIT methodology assumes the independence of all the messages exchanged on the factor graph of Fig. 5, by completely ignoring the presence of cycles. Fig. 9 compares the measured BER with the predicted BER obtained with the EXIT methodology, P_b. Since there is only a minor difference between the two curves, we conclude that the existence of cycles in the factor graph has only a small impact on the performances of the proposed method.

B. Comparison with existing message passing methods

We first consider the performances of iterative equalization for BPSK modulation.

Fig. 10 shows the bit error rate (BER) performances averaged over all the cells in the network, corresponding to channel A with parameter $a = 0.5$. $N_t = 5$ rounds of iterative equalization are performed for the IMS [10]-[11], BP [15]-[17] and GaBP equalization technique. Our experiments showed that augmenting the number of iterations did not improve the BER performances for any of the equalization method.

In order to assess the gap between actual and optimal performances, a single user lower bound on the BER is also computed by considering that the total energy collected at BS (i,j) originates from the user in cell (i,j). For BPSK, the expression of this lower bound is given by

$$P_{BPSK}(\sigma) = E\left[Q\left(\sqrt{\frac{2\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|h_{i,j}^{m,n}|^2}{\sigma^2}}\right)\right],$$

where the expectation is over all channel realizations. We observe that, at high SNR, IMS-based iterative equalization is quasi-optimal, while BP-based and GaBP-based iterative equalization suffer from a 0.9 dB and 1.4 dB loss in terms of power efficiency, respectively. Considering that GaBP equalization, as opposed to IMS and BP equalization, has a complexity which is not exponential in the size of the state-space, we conclude that GaBP-based iterative equalization trades a small loss in terms of power efficiency against a lower implementation cost.

Fig. 11 shows the same results for channel A with parameter $a = 0.75$. At high SNR, both BP-based and GaBP-based iterative equalization suffer from a 1.2 dB loss in terms of power efficiency, while IMS-based iterative equalization is again quasi-optimal.

C. Influence of the channel parameters in the Aktas et al. model

Fig. 12 illustrates the BER performance of the proposed iterative GaBP equalizer with BPSK modulation on channel A, for different values of the channel parameter a. We observe that for a fixed SNR, the BER increases with a, as a result of a higher spatial diversity advantage provided by cooperative base station processing. Also, the performance loss with respect to the single user bound decreases with a. The results for channel B, with different values of the channel parameters a and b, are shown in Fig 13.

Results for a more realistic cellular network model, including path loss, shadowing and multipath fading models, are provided in Sec. V-F.

D. Higher order modulations

We study the performances of GaBP-based iterative equalization using QPSK and 8-PSK with Gray labeling.

For QPSK and 8-PSK, the expression of the single user lower bound is obtained as

$$P_{QPSK}(\sigma) = E\left[Q\left(\sqrt{\frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|h_{i,j}^{m,n}|^2}{\sigma^2}}\right)\right],$$

$$P_{8-PSK}(\sigma) = E\left[\frac{2}{3}Q\left(\sqrt{\frac{2\sin(\pi/8)^2\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|h_{i,j}^{m,n}|^2}{\sigma^2}}\right)\right].$$

Fig. 14 and show the BER performances of GaBP-based iterative equalization, on channel A with parameter $a = 0.5$, for QPSK modulation. The suboptimality of GaBP-based iterative equalization vs. the single user lower bound leads to a power efficiency loss of 3 dB for QPSK modulation, after $N_t = 6$ iterations.

Fig. 15 shows the same results for 8-PSK modulation. The suboptimality of GaBP-based iterative equalization vs. the single user lower bound leads to a power efficiency loss of 2.6 dB for 8-PSK modulation, after $N_t = 10$ iterations.

No further improvements could be observed by augmenting the number of iterations. IMS and BP-based iterative equalization have not been implemented, due to their prohibitive complexity when higher order modulations are used.

E. Long memory channel example

In order to further illustrate the complexity advantage of the proposed algorithm over existing IMS and BP-based iterative equalization, we now consider channel B with parameters $a = 0.5$ and $b = 0.1$, corresponding to an hexagonal network with intercell interference limited to the 18 nearest neighboring cells. For this experiment, we use again BPSK modulation for simplicity. In this case, IMS equalization would be a formidable task, since each

BCJR equalizer would have to work on a trellis with 2^{20} states (i.e. more than one million possible discrete states). For similar reasons, BP equalization would also be intractable.

We compare two tractable approaches with reasonable computational complexity

- the proposed GaBP-based iterative equalization with full 2D channel memory
- an IMS algorithm with an embedded reduced-state BCJR with truncated 2D channel memory inspired by [42], where at the current iteration, an estimate of the residual 2D ISI and its variance are obtained by computing soft-decisions based on the LLRs computed at the previous iteration.

Fig. 16-17 show the BER performances after $N_t = 5$ rounds of iterative equalization on channel B. The proposed GaBP-based method with full 5×5 channel memory outperforms the reduced-state IMS with truncated 3×3 memory.

F. Realistic propagation modeling in a cellular network

In Sec. II-A, channel A and channel B were introduced as simple propagation models for the uplink of hexagonal cellular networks. We now consider a realistic propagation model including the effect of path loss, shadowing and multipath fading [43], [17]. Assuming full frequency reuse, each BS receives the desired signal from the user in its cell and intercell interference from the users in all other cells. The channel coefficient from the user in cell (i,j) to the BS in cell (k,l) can be modeled as

$$h_{(i,j)\to(k,l)} = \sqrt{P_{(i,j)\to(k,l)}}\alpha_{(i,j)\to(k,l)}, \tag{23}$$

where $P_{(i,j)\to(k,l)}$ (resp.$\alpha_{(i,j)\to(k,l)}$) denotes the power (resp. the multipath fading coefficient) of the signal received at the BS in cell (k,l) from the user in cell (i,j). Assuming flat Rayleigh fading, the coefficients $\alpha_{(i,j)\to(k,l)}$ are drawn independently from a complex circular Gaussian distribution with zero mean and unit variance [43]. In order to model the propagation path loss and shadowing due to large obstacles [43], we let

$$P_{(i,j)\to(k,l)} = K\frac{P_{TX}^{(i,j)}}{d_{(i,j)\to(k,l)}^{\gamma}}10^{\frac{\Psi_{(i,j)\to(k,l)}}{10}},$$

where K is a constant, $P_{TX}^{(i,j)}$ is the transmit power of the user in cell (i,j), $d_{(i,j)\to(k,l)}$ is the distance between the user in cell (i,j) and the BS in cell (k,l) and γ is the path loss exponent. The shadowing exponents $\Psi_{(i,j)\to(k,l)}$ are drawn independently from a real Gaussian distribution with zero mean and standard deviation σ_Ψ dB [43]. In our simulations, the distance between base stations is 1 km, $\gamma = 3.5$ and $\sigma_\Psi = 4$ dB. Assuming perfect power control of the user in each cell with its intended BS, we obtain the constraints

$$P_{(i,j)\to(i,j)} = \Gamma, \quad \forall(i,j),$$

where Γ is a constant, so that the SNR of a user with the intended BS is Γ/σ^2. These constraints are sufficient to specify entirely the value of all channel coefficients in (23).

Fig. 18 shows the BER performances after $N_t = 5$ rounds of the proposed iterative GaBP equalizer with BPSK modulation obtained via Monte Carlo simulations, where each user location is drawn uniformly at random in the hexagonal cell it belongs to. For a 2D channel memory truncated to $(M = 3, N = 3)$, very poor results are obtained. For a 2D channel memory truncated to $(M = 5, N = 5)$, at low and medium SNR values the BER approches the single user bound, while at high SNR where the truncated interference plays a dominant role, an error floor occurs.

VI. Conclusion

In this paper, we introduced a soft-output 2D equalization method, which is suitable for iterative equalization on 2D ISI channels. The proposed method processes each line of observations using Gaussian belief propagation (GaBP) on a suitable factor graph representation. Since the actual state-space is discrete, parameterizing the messages exchanged on the factor graph as Gaussian distributions is a crude approximation. Nevertheless, satisfactory performances were obtained for intercell interference cancellation for the uplink of wireless cellular networks with distributed BS processing.

The benefit of the proposed algorithm is twofold. Firstly, the GaBP equalizer has a complexity which is independent of the size of the constellation used for modulation. Secondly, the computational complexity of the

proposed GaBP-based equalization is much lower than its BCJR or BP-based counterparts for 2D channels with long memory.

Future work will consider GaBP-based equalization performed jointly with channel decoding and channel estimation.

APPENDIX A
NON-CAUSAL CHANNEL MODELS

Assume a rectangular $I \times J$ network of BSs. With the standard assumption that the multi-access interference is confined to adjacent cells, the received observation at the BS of cell (i, j) has the form

$$r_{i,j} = \sum_{k=-M_0}^{M_0} \sum_{l=-N_0}^{N_0} f_{i,j}^{k+M_0,l+N_0} b_{i-k,j-l} + w_{i,j}.$$

where $\{f_{m,n}^{i,j}\}$ are the coefficients of the 2D multi-access interference channel and $w_{i,j}$ is an additive white Gaussian noise (AWGN) term. As suggested in [11], by introducing the delay (M_0, N_0), we obtain

$$r_{i-M_0,j-N_0} = \sum_{k=-M_0}^{M_0} \sum_{l=-N_0}^{N_0} f_{i-M_0,j-N_0}^{k+M_0,l+N_0} b_{i-k-M_0,j-l-N_0} + w_{i-M_0,j-N_0}.$$

Using the change of variables $m = k + M_0$, $n = l + N_0$

$$r_{i-M_0,j-N_0} = \sum_{m=0}^{2M_0} \sum_{n=0}^{2N_0} f_{i-M_0,j-N_0}^{m,n} b_{i-m,j-n} + w_{i-M_0,j-N_0}.$$

Now let $M = 2M_0 + 1$, $N = 2N_0 + 1$, $y_{i,j} = r_{i-M_0,j-N_0}$, $h_{i,j}^{m,n} = f_{i-M_0,j-N_0}^{m,n}$ and $n_{i,j} = w_{i-M_0,j-N_0}$, the causal model in (1) is obtained.

This reasoning can easily be extended to a cellular network with hexagonal 2D cellular structure.

APPENDIX B
PROOF OF THE FACTORIZATION IN EQUATION (5)

The *a posteriori* pmf of the input vectors and the 2D ISI states corresponding to the i-th line can be written as

$$\begin{aligned} p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J} | y_{i,1:J}) &\propto p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J}) p(y_{i,1:J} | \mathbf{i}_{1:J}, \mathbf{x}_{1:J}) \\ &\propto p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J}) \prod_{k=1}^{J} p(y_{i,k} | \mathbf{x}_k). \end{aligned} \tag{24}$$

where the second equation follows from the fact that the observation noise term in (4), $n_{i,k}$, is white. According to Bayes rule, we also have

$$\begin{aligned} p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J}) &= p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J-1}) p(\mathbf{x}_J | \mathbf{i}_{1:J}, \mathbf{x}_{1:J-1}) \\ &= p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J-1}) p(\mathbf{x}_J | \mathbf{i}_J, \mathbf{x}_{J-1}). \end{aligned}$$

where the second equality follows from the first-order Markov assumption for the 2D ISI states in (4). Now, the reasoning applied at time instant $k = J$ can be applied recursively for all time instants $k = J - 1, \ldots, 1$, which leads to

$$p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J}) = p(\mathbf{i}_{1:J}, \mathbf{x}_0) \prod_{k=1}^{J} p(\mathbf{x}_k | \mathbf{i}_k, \mathbf{x}_{k-1}).$$

Now, assuming that the initial 2D ISI state $\mathbf{x}_0$ is independent from the sequence of input vectors $\mathbf{i}_{1:J}$, we have

$$p(\mathbf{i}_{1:J}, \mathbf{x}_{1:J}) = p(\mathbf{x}_0) \prod_{k=1}^{J} p(\mathbf{i}_k) p(\mathbf{x}_k | \mathbf{i}_k, \mathbf{x}_{k-1}), \tag{25}$$

where we have used the fact that the data symbols are i.i.d., thus the input vectors are also independent. Combining (24) and (25) yields the desired result.

APPENDIX C
PROOF OF EQUATION (10)

Since $\mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1})$ does not depend on $\mathbf{i}_k$, Eq. (9) can be rewritten as

$$\begin{aligned}\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k) &= \int\left(\int p(\mathbf{x}_k|\mathbf{i}_k,\mathbf{x}_{k-1})p(\mathbf{i}_k)d\mathbf{i}_k\right)\mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1})d\mathbf{x}_{k-1}\\ &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1})\mu_{\mathbf{x}_{k-1}\to f_k}(\mathbf{x}_{k-1})d\mathbf{x}_{k-1}.\end{aligned} \tag{26}$$

From Sec. III-A, we have $\mathbf{i}_k \sim \mathcal{N}(\mathbf{i}_k : \hat{\mathbf{i}}_k, \mathbf{Q}_k)$. Moreover, in Sec. II, we assumed that the symbols are independent, therefore $\mathbf{i}_k$ and $\mathbf{x}_{k-1}$ are independent. Then it follows from (4) that $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is a Gaussian distribution whose mean is calculated as

$$E[\mathbf{x}_k|\mathbf{x}_{k-1}] = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\hat{\mathbf{i}}_k$$

and whose covariance is calculated as

$$\begin{aligned}&E[(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1} - \mathbf{G}\hat{\mathbf{i}}_k)(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1} - \mathbf{G}\hat{\mathbf{i}}_k)^H|\mathbf{x}_{k-1}]\\ &= E[(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1} - \mathbf{G}\mathbf{i}_k + \mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k))(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1} - \mathbf{G}\mathbf{i}_k + \mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k))^H|\mathbf{x}_{k-1}]\\ &= E[(\mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k))(\mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k))^H|\mathbf{x}_{k-1}]\\ &= \mathbf{G}\mathbf{Q}_k\mathbf{G}^H.\end{aligned}$$

Combining this result with (8), (26) becomes

$$\mu_{f_k\to\mathbf{x}_k}(\mathbf{x}_k) = \int \mathcal{N}(\mathbf{x}_k : \mathbf{F}\mathbf{x}_{k-1} + \mathbf{G}\hat{\mathbf{i}}_k, \mathbf{G}\mathbf{Q}_k\mathbf{G}^H)\mathcal{N}(\mathbf{x}_{k-1} : \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})d\mathbf{x}_{k-1}.$$

We recognize the prediction integral of the well-known Kalman filter [30], which yields the desired result.

APPENDIX D
PROOF OF EQUATION (15)

If the approximations

$$\begin{cases}\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0 : \hat{\mathbf{x}}_0, \mathbf{P}_0)\\ \mathbf{i}_k \sim \mathcal{N}(\mathbf{i}_k : \hat{\mathbf{i}}_k, \mathbf{Q}_k), \quad k = 1,\dots,J\end{cases}$$

introduced in Sec. III-A hold, then it follows from (4) that $\mathbf{x}_k$, $k = 1,\dots,J$ is also Gaussian distributed. Let $\hat{\mathbf{x}}_k$ and $\mathbf{\Pi}_k$ denote the *a priori* mean and covariance of $\mathbf{x}_k$, respectively. We have

$$\begin{aligned}\hat{\mathbf{x}}_k &= E[\mathbf{x}_k] = \mathbf{F}E[\mathbf{x}_{k-1}] + \mathbf{G}E[\mathbf{i}_k]\\ &= \mathbf{F}\hat{\mathbf{x}}_{k-1} + \mathbf{G}\hat{\mathbf{i}}_k\end{aligned}$$

and

$$\begin{aligned}\mathbf{\Pi}_k &= E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^H]\\ &= E\{[\mathbf{F}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}) + \mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k)][\mathbf{F}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}) + \mathbf{G}(\mathbf{i}_k - \hat{\mathbf{i}}_k)]^H\}\\ &= \mathbf{F}\mathbf{\Pi}_{k-1}\mathbf{F}^H + \mathbf{G}\mathbf{Q}_k\mathbf{G}^H,\end{aligned}$$

where the last equality holds because $\mathbf{i}_k$ and $\mathbf{x}_{k-1}$ are independent (see Sec. II).

Then, a straightforward application of proposition 7 in [32] gives a closed form expression of the backward dynamics of the 2D ISI states as

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}) = \mathcal{N}(\mathbf{x}_k : \tilde{\mathbf{F}}_{k+1}\mathbf{x}_{k+1} + \mathbf{c}_{k+1}, \tilde{\mathbf{Q}}_{k+1}),$$

where

$$\begin{cases}\tilde{\mathbf{F}}_{k+1} = \mathbf{\Pi}_k\mathbf{F}^H\mathbf{\Pi}_{k+1}^{-1}\\ \mathbf{c}_{k+1} = \hat{\mathbf{x}}_k - \tilde{\mathbf{F}}_{k+1}\hat{\mathbf{x}}_{k+1}\\ \tilde{\mathbf{Q}}_{k+1} = (\mathbf{I}_{MN} - \tilde{\mathbf{F}}_{k+1}\mathbf{F})\mathbf{\Pi}_k.\end{cases}$$

REFERENCES

[1] W. Coene, "Two-dimensional optical storage," *Technical Digest of Optical Data Storage (ODS) Conference*, pp. 90-92, Vancouver, Canada, 2003.

[2] J. Ashley, M.-P. Bernal, G.W. Burr, H. Coufal, H. Guenther, J.A. Hoffnagle, C.M. Jefferson, B. Marcus, R.M. Macfarlane, R.M. Shelby and G.T. Sincerbox, "Holographic data storage," *IBM J. Res. Develop.*, vol. 44, no. 3, pp. 341-368, May 2000.

[3] Z. Zhang, T.M. Duman and E.M. Kurtas, "Achievable information rates and coding for MIMO systems over ISI channels and frequency-selective fading channels," *IEEE Trans. Comm.*, pp. 1698-1710, vol. 52, no. 10, Oct. 2004.

[4] A.D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inform. Theory*, pp. 1713-1727, vol. 40, no. 6, Nov. 1994.

[5] G.D. Forney, "The Viterbi algorithm," *Proc. IEEE*, pp. 268–278, vol. 61, no. 3, Mar. 1973.

[6] X. Chen and K.M. Chugg, "Near-optimal data detection for two-dimensional ISI/AWGN channels using concatenated modeling and iterative algorithms," *Proc. ICC 98*, pp. 952-956, Atlanta, USA, 1998.

[7] K.M. Chugg, X. Chen and M.A. Neifeld, "Two-dimensional equalization in coherent and incoherent page-oriented optical memory," *J. Opt. Soc. Am. A*, pp. 549-562, vol. 46, no. 3, Mar. 1999.

[8] J.K. Nelson, A.C. Singer and U. Madhow, "Multi-directional decision feedback for 2D equalization," *Proc. ICASSP 04*, pp. 921-924, Montreal, Canada, May 2004.

[9] J.F. Heanue, K. Gurkan and L. Hesselink, "Signal detection for page-access optical memories with intersymbol interference," *Applied Optics*, vol. 35, no. 14, pp. 2431-2438, May 1996.

[10] M. Marrow and J.K. Wolf, "Detection of 2-dimensional signals in the presence of ISI and noise," *Proc. ISITA 04*, pp. 891-894, Parma, Italy, Oct. 2004.

[11] J.B. Soriaga, P.H. Siegel, J.K. Wolf and M. Marrow, "On achievable rates of multistage decoding on two-dimensional ISI channels," *Proc. ISIT 05*, pp. 1348-1352, Adelaide, Australia, Sept. 2005.

[12] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, Vol. 20, pp. 284-287, March 1974.

[13] T. Cheng, B.J. Belzer and K. Sivakumar, "Row-column soft-decision feedback algorithm for two-dimensional intersymbol interference," *IEEE Sig. Proc. Lett.*, vol. 14, no. 7, pp. 433-436, July 2007.

[14] F.R. Kschischang, B.J. 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.

[15] N. Singla and J.A. O'Sullivan, "Joint equalization and decoding for nonlinear two-dimensional intersymbol interference channels," *Proc. ISIT 05*, pp. 1353-1357, Adelaide, Australia, Sept. 2005.

[16] E. Aktas, J. Evans and S. Hanly, "Distributed decoding in a cellular multiple-access channel," *IEEE Trans. Wireless Comm.*, vol. 7, no. 1, pp. 241-250, Jan. 2008.

[17] S. Bavarian and J.K. Cavers, "Reduced-complexity belief propagation for system-wide MUD in the uplink of cellular networks," *JSAC*, vol. 26, no. 3, pp. 541-549, Apr. 2008.

[18] J. Pearl, *Probabilistic reasoning in intelligent systems*, San Francisco, CA: Morgan Kaufmann, 1988.

[19] 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.

[20] O. Shental, N. Shental, S. Shamai, I. Kanter, A.J. Weiss and Y. Weiss, "Discrete-input two-dimensional Gaussian channels with memory: estimation and information rates via graphical models and statistical mechanics," *IEEE Trans. Inform. Theory*, vol. 54, no. 4, pp. 1500-1513, Apr. 2008.

[21] T.P. Minka, "Expectation propagation for approximate Bayesian inference," *Proc. UAI 01*, pp. 362-369, Seattle, USA, 2001.

[22] Y. Weiss and W.T. Freeman, "Correctness of belief propagation in Gaussian graphical models of arbitrary topology," *Neural Computation*, vol. 13, no. 10, pp. 2173-2200, 2001.

[23] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Sig. Proc. Mag.*, vol. 21, no. 1, pp. 28-41, Jan. 2004.

[24] J. Hu, H.-A. Loeliger, J. Dauwels and F. Kschischang, "A general computation rule for lossy summaries/messages with examples from equalization," *Proc. 44th Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sept. 2006.

[25] D.M. Malioutov, J.K. Johnson and A.S. Willsky, "Walk-sums and belief propagation in Gaussian graphical models," *The Journal of Machine Learning Research*, vol. 7, pp. 2031-2064, Oct. 2006.

[26] D. Bickson and D. Dolev, O. Shental, P.H. Siegel and J.K. Wolf, "Gaussian belief propagation based multiuser detection," *Proc. ISIT 08*, pp. 1878-1882, Toronto, Canada, Jul. 2008.

[27] F. Lehmann, "Turbo equalization of two-dimensional intersymbol interference channels using Gaussian belief propagation," *Proc. ICASSP 2010*, pp. 3182-3185, Dallas, TX, USA, March 2010.

[28] Q. Guo and L. Ping, "LMMSE turbo equalization based on factor graphs," *JSAC*, vol. 26, no. 2, pp. 311-319, Feb. 2008.

[29] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping and F.R. Kschischang, "The factor graph approach to model-based signal processing," *Proc. IEEE*, pp. 1295–1322, vol. 95, no. 6, June 2007.

[30] B.D.O. Anderson and J.B. Moore, *Optimal filtering*, Englewood Cliffs, N.J.: Prentice Hall, 1979.

[31] D.C. Fraser and J.E. Potter, "The optimum linear smoother as a combination of two optimum linear filters," *IEEE Trans. Automat. Contr.*, vol. 14, no. 4, pp. 387-390, Aug. 1969.

[32] B. Ait-El-Fquih and F. Desbouvries, "On Bayesian fixed-interval smoothing algorithms," *IEEE Trans. Automat. Contr.*, vol. 53, no. 10, pp. 2437-2442, Nov. 2008.

[33] C. Laot, R. Le Bidan and D. Leroux, "Low-complexity MMSE turbo equalization: a possible solution for EDGE," *IEEE Trans. Wireless Comm.*, vol. 4, no. 3, pp. 965-974, May 2005.

[34] L. Papke, P. Robertson and E. Villebrun, "Improved decoding with the SOVA in a parallel concatenated (Turbo-code) scheme," *Proc. ICC 96*, pp. 102-106, Dallas, TX, USA, June 1996.

[35] R.M. Pyndiah, ”Near optimum decoding of product codes: block turbo codes,” *IEEE Trans. Comm.*, vol. 46, no. 8, pp. 1003-1010, Aug. 1998.
[36] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” *IEEE Trans. Comm.*, vol. 49, no. 10, pp. 1727-1737, Oct. 2001.
[37] J. Kliewer, S.X. Ng and L. Hanzo, “Efficient computation of EXIT functions for nonbinary iterative decoding,” *IEEE Trans. Comm.*, vol. 54, no. 12, pp. 2133-2136, Dec. 2006.
[38] R. Zhang and L. Hanzo, “Cochannel interference mitigation - active and passive techniques,” *IEEE Vehicular Technology Mag.*, pp. 31-39, Dec. 2010.
[39] L.E. Baum, T. Petrie, G. Soules and N. Weiss, ”A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains,” Ann. Math. Statist. , vol. 41, no. 1, pp. 164-171, 1970.
[40] W. Bobillet, R. Diversi, E. Grivel, R. Guidorzi, M. Najim and U. Soverini, “Speech enhancement combining optimal smoothing and errors-in-variables identification of noisy AR processes,” *IEEE Trans. Sig. Proc.*, vol. 55, no. 12, pp. 5564-5578, Dec. 2007.
[41] A.H.J. Immink, W.M.J. Coene, A.M. van der Lee, C. Busch, A.P. Hekstra, J.W.M. Bergmans, J. Riani, S.J.L. v. Beneden and T. Conway, “Signal processing and coding for two-dimensional optical storage,” *Proc. Globecom 03*, pp. 3904-3908, San Francisco, USA, Dec. 2003.
[42] B. Baccarelli, A. Fasano and A. Zucchi, “A reduced-state soft-statistics-based MAP/DF equalizer for data transmission over long ISI channels,” *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1441-1446, Sept. 2000.
[43] G.L. Stüber, *Principles of Mobile Communications.* Kluver Academic Publishers, Norwell, Massachusetts, 1999.

Frederic Lehmann received the E.E. degree and the M.S.E.E. degree from ENSERG, France, in 1998. In 2002, he received the PhD in Electrical Engineering from the National Polytechnical Institute, Grenoble (INPG), France. He worked as a Research Engineer with STMicroelectronics from 1999 to 2002. From 2003 to 2004 he was a Post-doctoral Researcher at LAAS (Laboratory for Analysis and Architecture of Systems), CNRS, Toulouse, France. Currently, he is an Assistant Professor at Institut TELECOM, Telecom SudParis, Evry, France. His main research interests are in the area of communication theory, non-linear signal processing and statistical image processing.

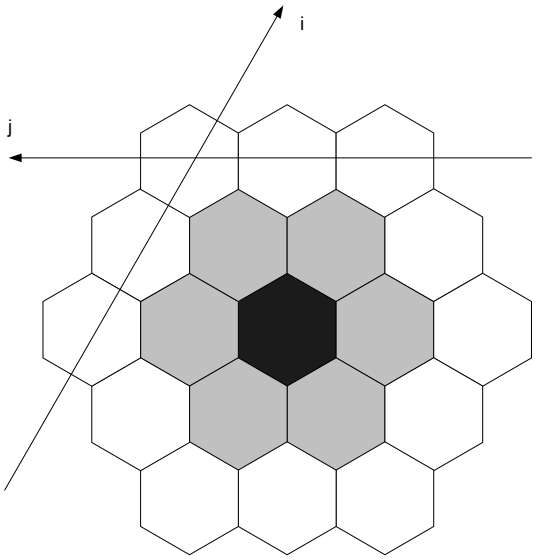

Fig. 1. Hexagonal cellular network structure. The cell in black is depicted along with its 6 nearest neighboring cells (gray cells) and its 18 nearest neighboring cells (gray and white cells)

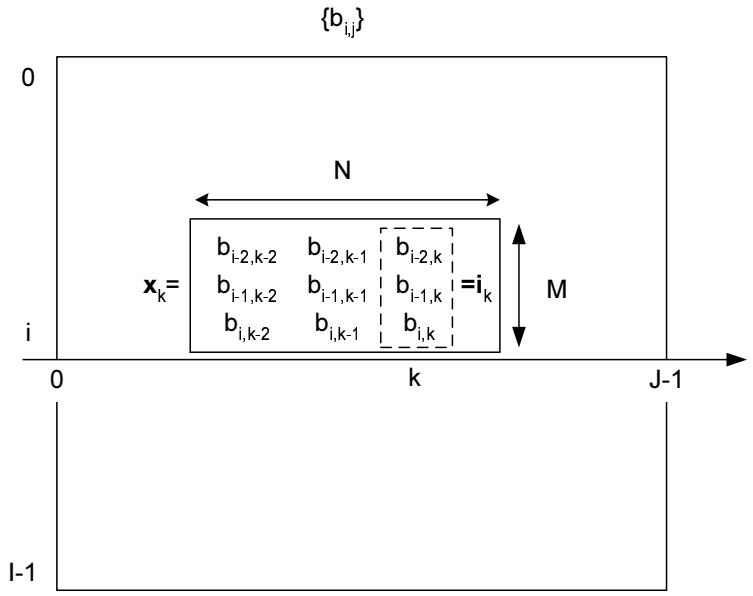

Fig. 2. State space representation corresponding to the i-th line of observations (here $M = 3$, $N = 3$).

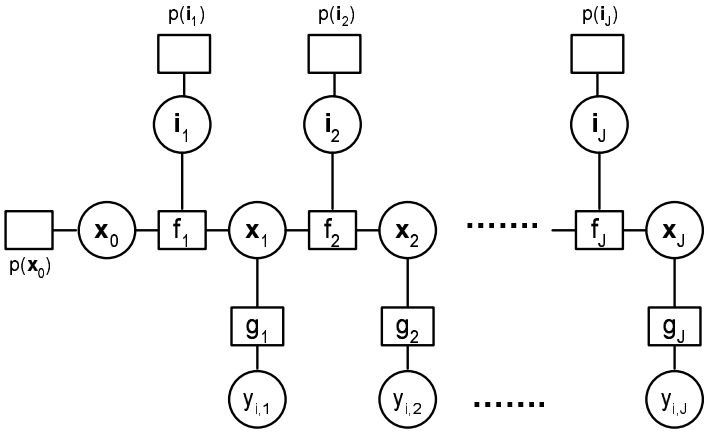

Fig. 3. Factor graph corresponding to the i-th line of observations. Variable nodes are represented by circles and function nodes by squares.

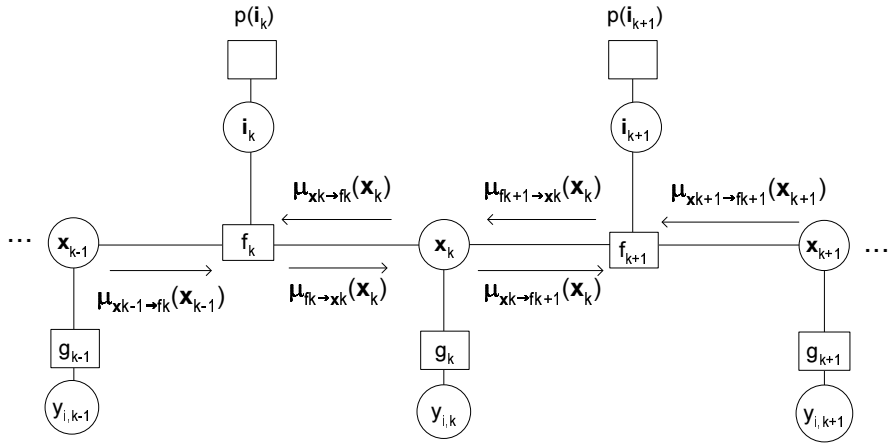

Fig. 4. Messages exchanged on a portion of the factor graph corresponding to the i-th line of observations.

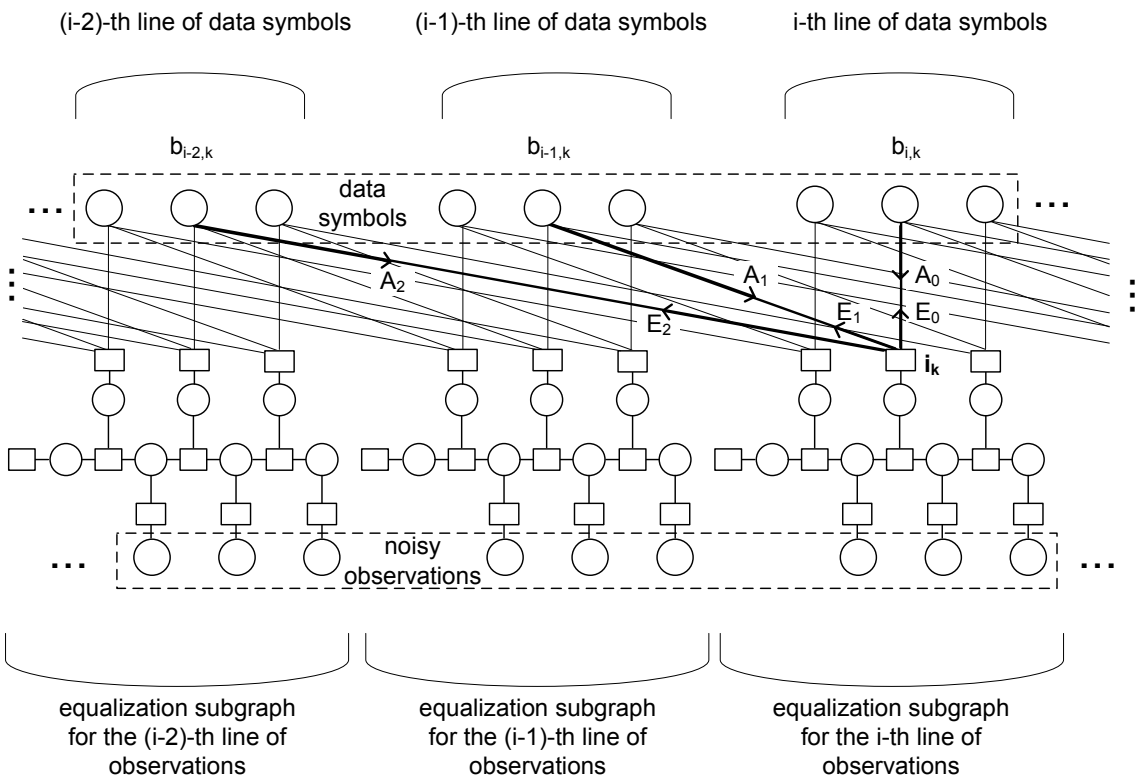

Fig. 5. Complete factor graph for channel A (i.e. $M = 3$). Data symbols and noisy observations are linewise raster scanned. A_m (resp. E_m), $0 \leq m \leq M-1$, are the *a priori* (resp. *extrinsic*) values sent by a data symbol node (GaBP equalization subgraph) to a GaBP equalization subgraph (resp. data symbol node).)

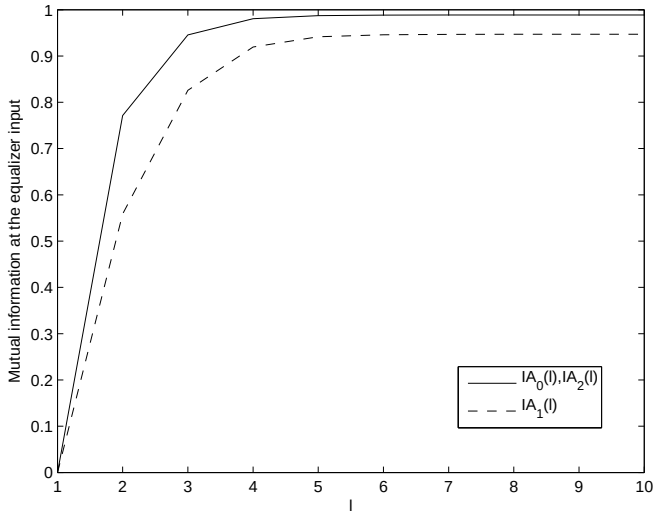

Fig. 6. GaBP-based iterative equalization for channel A with $a = 0.5$ and BPSK: evolution of the input mutual information as a function of the iteration index l.

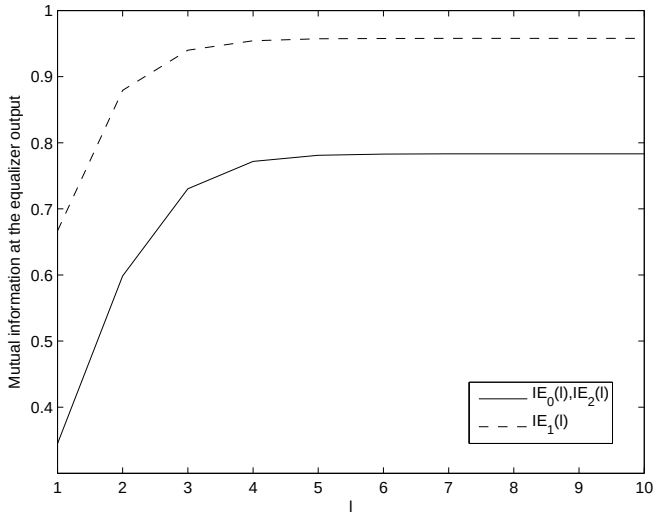

Fig. 7. GaBP-based iterative equalization for channel A with $a = 0.5$ and BPSK: evolution of the output mutual information as a function of the iteration index l.

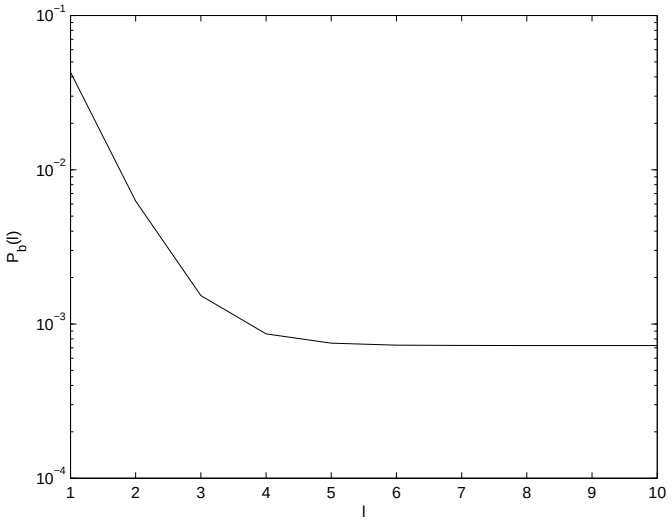

Fig. 8. GaBP-based iterative equalization for channel A with $a = 0.5$ and BPSK: evolution of the predicted BER, P_b, as a function of the iteration index l.

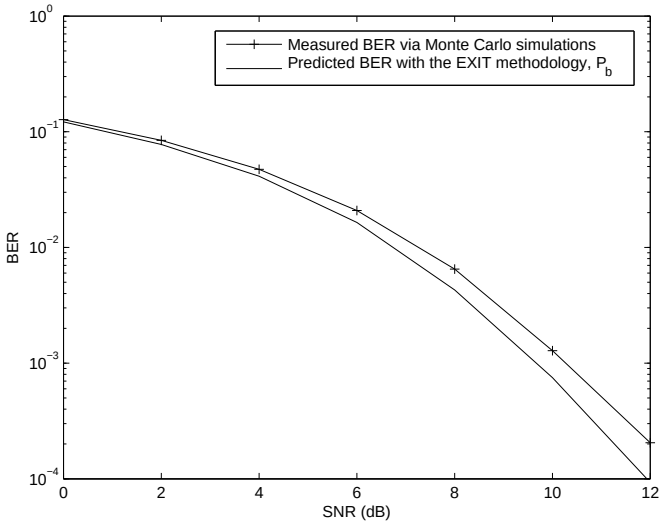

Fig. 9. GaBP-based iterative equalization for channel A with $a = 0.5$ and BPSK: measured vs. predicted BER.

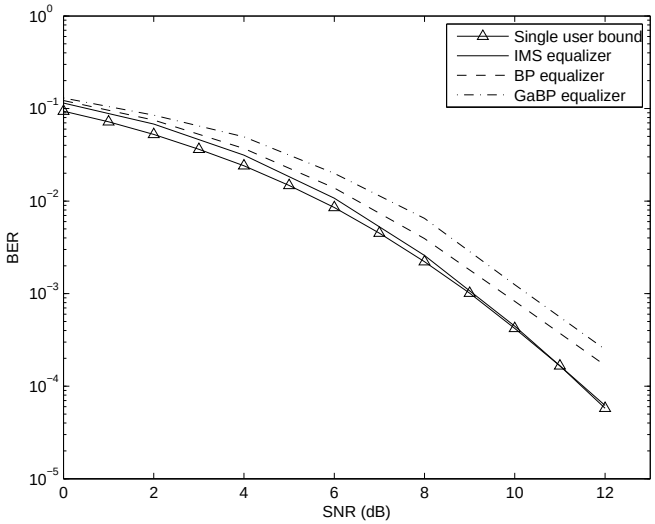

Fig. 10. Iterative equalization for channel A with $a = 0.5$ and BPSK ($N_t = 5$ iterations).

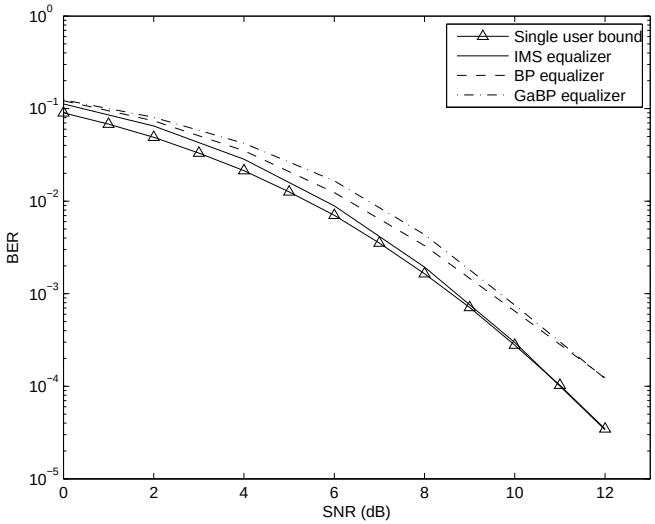

Fig. 11. Iterative equalization for channel A with $a = 0.75$ and BPSK ($N_t = 5$ iterations).

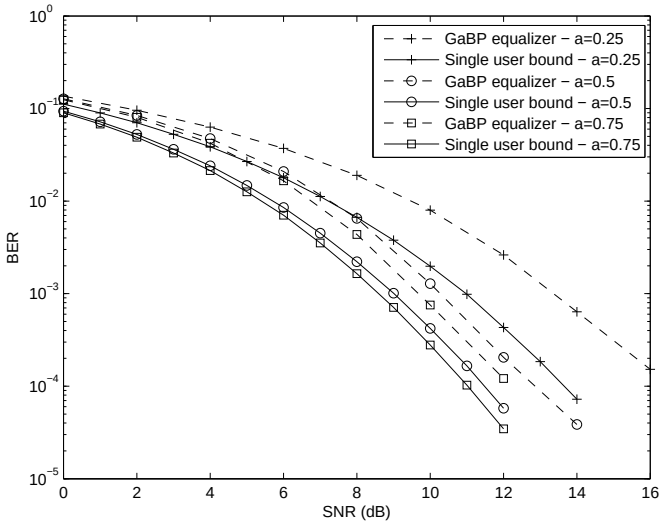

Fig. 12. BER of iterative GaBP equalization ($N_t = 5$ iterations) for BPSK modulation on channel A, for different values of parameter a.

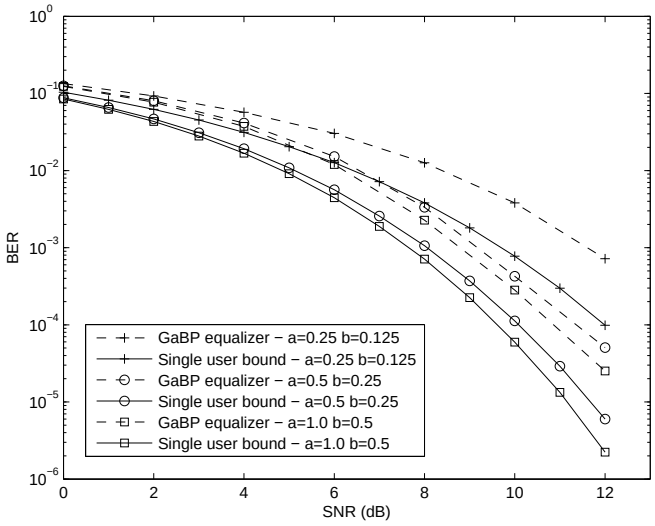

Fig. 13. BER of iterative GaBP equalization ($N_t = 5$ iterations) for BPSK modulation on channel B, for different values of parameters a and b.

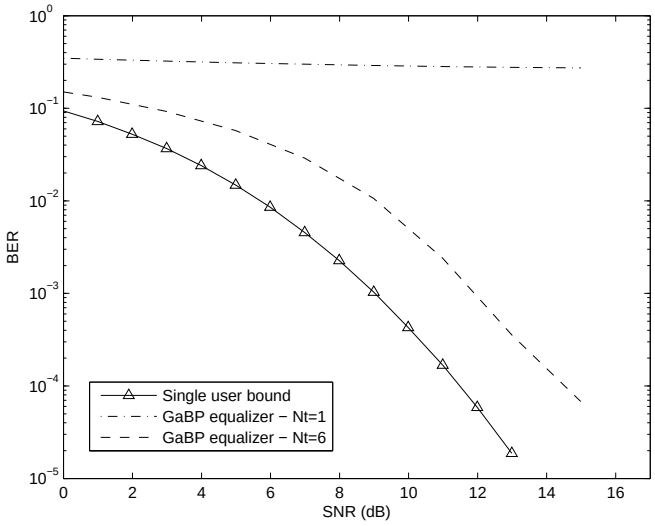

Fig. 14. Iterative equalization for channel A with $a = 0.5$ and QPSK.

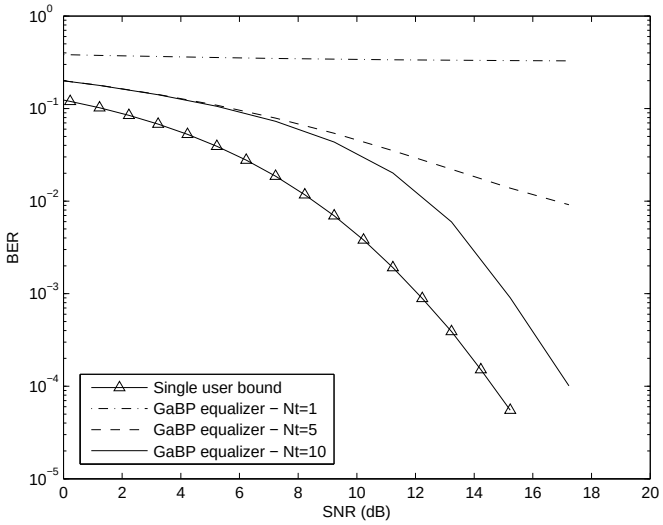

Fig. 15. Iterative equalization for channel A with $a = 0.5$ and 8-PSK.

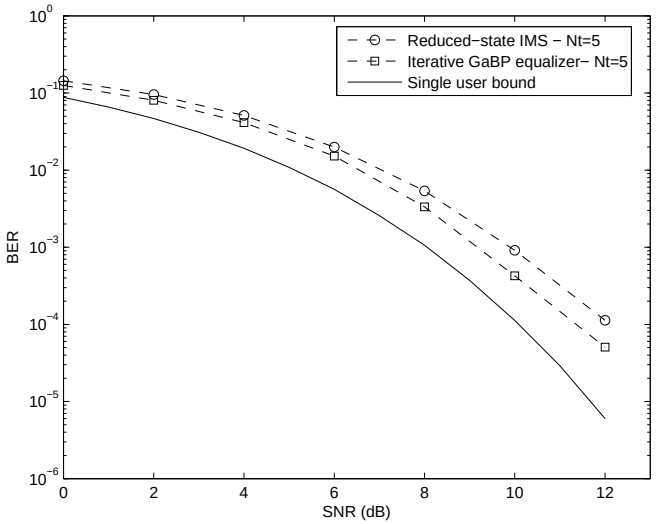

Fig. 16. Comparison of iterative equalization for channel B with $a = 0.5$, $b = 0.25$ and BPSK ($N_t = 5$ iterations) using the proposed GaBP method with full 5×5 channel memory and the reduced-state IMS with truncated 3×3 memory.

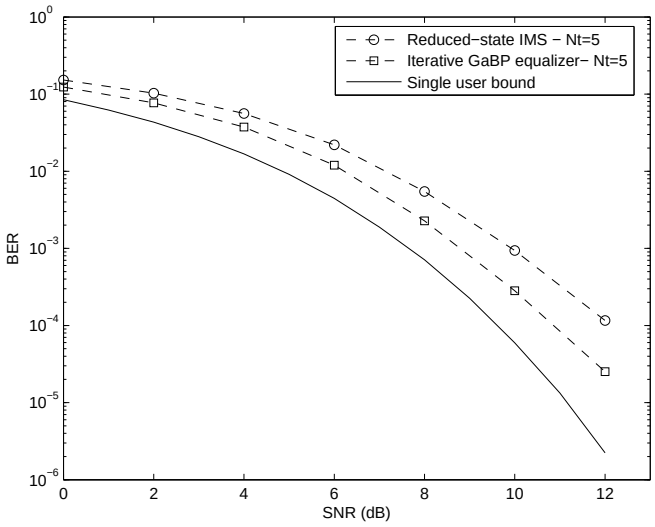

Fig. 17. Comparison of iterative equalization for channel B with $a = 1.0$, $b = 0.5$ and BPSK ($N_t = 5$ iterations) using the proposed GaBP method with full 5×5 channel memory and the reduced-state IMS with truncated 3×3 memory.

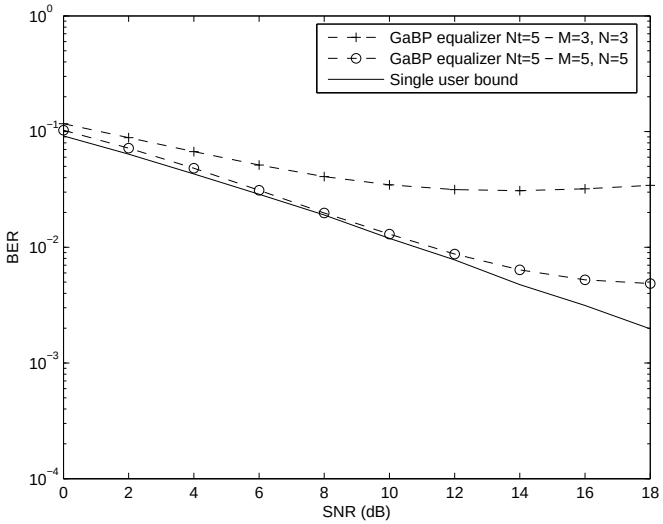

Fig. 18. Iterative equalization for a realistic propagation model and BPSK ($N_t = 5$ iterations) using the proposed GaBP method with 3×3 and 5×5 channel memory.