Channel Estimation Using Gaussian Approximation in a Factor Graph for QAM Modulation

Yang Liu[*][†], Loïc Brunel[†], and Joseph J. Boutros[‡]

[*]ENST, 46 rue Barrault, 75013 Paris, France - yliu@enst.fr
[†]Mitsubishi Electric ITE-TCL, 1 allée de Beaulieu, 35700 Rennes, France - brunel@tcl.ite.mee.com
[‡]Texas A&M University at Qatar, PO Box 23874, Doha, Qatar - boutros@tamu.edu

Abstract—**Joint channel estimation and decoding using belief propagation on factor graphs requires the quantization of probability densities since continuous parameters are involved. We propose to replace these densities by standard messages where the channel estimate is accurately modeled as a Gaussian mixture. Upward messages include symbol extrinsic information and downward messages carry a mean and a variance for the Gaussian modeled channel estimate. Such unquantized message propagation leads to a complexity reduction and a performance improvement. For QAM modulated symbols, the proposed belief propagation almost achieves the performance of Expectation-Maximization under good initialization and surpasses it under bad initialization.**

I. INTRODUCTION

Propagating messages in a suitable factor graph [1] is a systematic tool for deriving iterative algorithms. Among various receiver issues solved using the belief propagation algorithm (BP), also called sum-product algorithm [2], we can cite decoding, channel estimation, synchronization and detection [3]. [4] presents a BP handling continuous variables, in which canonical distributions are used for quantizing probability distributions, in order to propagate discrete probability distributions. However, the degree of quantization has a strong impact on estimation accuracy and performance. The expectation-maximization algorithm (EM) [5] [6], directly handling continuous variables through a posteriori probabilities (APPs) of transmitted symbols, achieves a better performance-complexity trade-off. Even adapting the quantization in each iteration of BP, as proposed in [7] and [8], does not fill the complexity gap between EM and BP.

Instead of relying on quantization, we propose here to model probability distributions as mixtures of Gaussian distributions. It allows for estimation improvement and complexity reduction simultaneously. High order QAM modulation is addressed here, so as to show that the Gaussian approximation principle also applies with a multiplicity of modulation symbols.

The paper is structured as follows. Section II explains how the transmission system is modeled using a factor graph and how BP is applied. Section III presents the approximation of the distribution of channel estimates in BP by a mixture of Gaussian distributions. In Section IV, APPs are computed from the approximated distribution. Channel estimation with continuous upward messages in the factor graph is presented in Section V. The paper ends with simulation results in Section VI.

In the sequel, messages that are not based on quantized densities will be referred to as continuous messages.

II. SYSTEM MODEL

We consider the transmission of N encoded M-QAM symbols x_k over a quasi-static single-input single-output (SISO) channel. Figure 1 depicts the encoding, interleaving and mapping of the information binary sequence b_i. The obtained M-QAM symbols x_k experience a single-tap channel, i.e., they are multiplied by a complex Gaussian channel coefficient $h \sim \mathcal{CN}(0,1)$ before addition of a complex Gaussian noise $n_k \sim \mathcal{CN}(0, 2\sigma_n^2)$. The receiver performs joint channel estimation and decoding based on the channel outputs y_k in order to obtain the estimated information sequence $\hat{b}_i$. The system model is described by

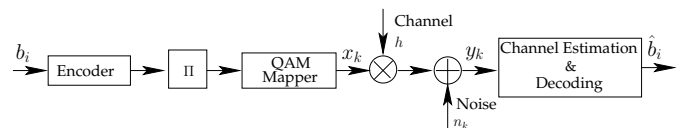$$y_k = h\,x_k + n_k, \qquad 0 \le k \le N-1. \tag{1}$$

Fig. 1.   System model.

The corresponding factor graph is built as in [4]. Figure 2 details the upward messages, where g is a quantized estimate of h and $p(g)$ represents the quantized distribution. In node CODE, a forward-backward algorithm computes the extrinsic information for each coded bit. Taken deinterleaving into account, the extrinsic information is propagated to nodes x_k, where the extrinsic information ξ_k for each M-QAM symbol x_k is computed. From each node f_k to node $p(g)$, a discrete distribution $\mu_{f_k \to p(g)}$ of the quantized estimate of h is computed and propagated based on a marginalization of the likelihood $p(y_k|x_k, g)$ with respect to the transmitted symbol x_k. All $\mu_{f_k \to p(g)}$ distributions are multiplied to build a single common discrete distribution $p(g)$. Figure 3 details the downward messages. The common distribution $p(g)$ is propagated down to all nodes f_k. The APP of each M-QAM symbol x_k is computed based on this discrete distribution, marginalizing the likelihood $p(y_k|x_k, g)$ with respect to g. The APP of each coded bit is then computed from the APP
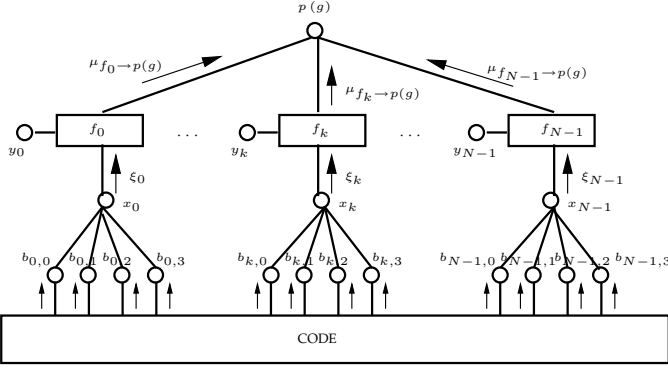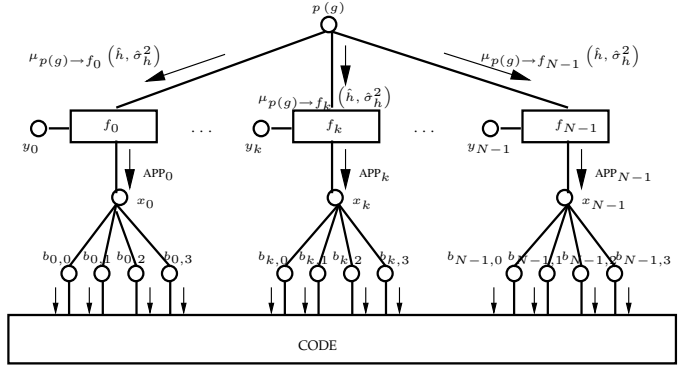
Fig. 2. Factor graph for upward message.

Fig. 3. Factor graph for downward message.

of the M-QAM symbol it belongs to. The whole process of propagating upward and downward messages is then iterated.

III. DISTRIBUTION OF CHANNEL ESTIMATE

In the iterative receiver, initial estimate is obtained from known pilots and subsequent estimates from data symbols. Thus, the distribution of channel estimate will differ depending on the iteration.

A. Estimation based on pilots

L_p pilots x_{pilot} are included in the transmitted sequence. From the L_p messages provided by $f_{\text{pilot},k}$ nodes corresponding to pilots, we get the discrete distribution of g [2]:

$$p(g) = \prod_{k=0}^{L_p-1} \mu_{f_{\text{pilot},k}\to p(g)}$$

$$\propto \sum_{i=0}^{q-1} \delta(g - g_i) \exp\left(-\frac{1}{2\sigma_n^2} \sum_{k=0}^{L_p-1} |y_{\text{pilot},k} - x_{\text{pilot},k} g_i|^2\right),$$

$$(2)$$

where

$$y_{\text{pilot},k} = h\, x_{\text{pilot},k} + n_{\text{pilot},k}, \qquad (3)$$

$\{g_i\}$ is a quantization codebook of size q for channel estimate's probability density function (pdf) and $\delta(\cdot)$ denotes the Dirac delta function. For $L_p \gg 1$, we have

$$\sum_{k=0}^{L_p-1} |y_{\text{pilot},k} - x_{\text{pilot},k} g_i|^2 \approx L_p \mathcal{E}_{\text{pilot}} |g_i - h|^2 + 2L_p\sigma_n^2, \quad (4)$$

where $\mathcal{E}_{\text{pilot}} = |x_{\text{pilot},k}|^2$ represents the pilot energy. Surprisingly, this rough approximation for small L_p also yields a good performance. Substituting (4) into (2), we obtain

$$p(g) \propto \sum_{i=0}^{q-1} \delta(g - g_i) \exp\left(-\frac{L_p \mathcal{E}_{\text{pilot}}}{2\sigma_n^2} |g_i - h|^2\right). \quad (5)$$

Hence, $p(g)$ can be approximated as a Gaussian distribution $\mathcal{CN}(h, \dfrac{2\sigma_n^2}{L_p \mathcal{E}_{\text{pilot}}})$.

B. Estimation based on data

Let $S = \{s_m, \ 0 \le m \le M-1\}$ represent the constellation of M-QAM and $\xi_{k,m}$ be the extrinsic information for $x_k = s_m$. The message from f_k to $p(g)$ can be expressed as [2]:

$$\mu_{f_k\to p(g)} \propto$$
$$\sum_{i=0}^{q-1} \delta(g - g_i) \sum_{m=0}^{M-1} \exp\left(-\frac{|y_k - s_m g_i|^2}{2\sigma_n^2}\right) \xi_{k,m}. \quad (6)$$

Using (1), we have

$$\mu_{f_k\to p(g)} \propto$$
$$\sum_{i=0}^{q-1} \delta(g - g_i) \sum_{m=0}^{M-1} \exp\left(-\frac{|s_m g_i - h\, x_k - n_k|^2}{2\sigma_n^2}\right) \xi_{k,m}. \quad (7)$$

Hence, the discrete distribution of g is expressed as (8), where s_k^j is the value of symbol x_k in sequence j and ξ_k^j is the extrinsic probability that x_k equals s_k^j. For $N \gg 1$, we get

$$\sum_{k=0}^{N-1} |s_k^j g_i - h\, x_k - n_k|^2$$

$$\approx \sum_{k=0}^{N-1} \left(|x_k h|^2 + |s_k^j g_i|^2 + 2\sigma_n^2 - s_k^{j*} x_k h g_i^* - s_k^j x_k^* h^* g_i\right). \quad (9)$$

In (9), considering the items $s_k^{j*} x_k h g_i^*$ and $s_k^j x_k^* h^* g_i$, the summation can be rewritten into five sub-summations where $s_k^j = x_k$, $s_k^j = -x_k$, $s_k^j = j x_k$, $s_k^j = -j x_k$ and s_k^j equals other values. We get (10), where U_1^j (resp. U_2^j, U_3^j and U_4^j) is the number of items with $s_k^j = x_k$ (resp. $s_k^j = -x_k$, $s_k^j = j x_k$ and $s_k^j = -j x_k$) and V represents the number of

$$p(g) = \prod_{k=0}^{N-1} \mu_{f_k\to p(g)} \propto \sum_{i=0}^{q-1} \delta(g - g_i) \sum_{j=1}^{M^N} \left\{ \exp\left(-\frac{1}{2\sigma_n^2} \sum_{k=0}^{N-1} |s_k^j g_i - h\, x_k - n_k|^2\right) \prod_{k=0}^{N-1} \xi_k^j \right\}. \quad (8)$$

2

all other items. $\mathcal{E}_{\text{av}}$ is the average energy of M-QAM symbols. Substituting (10) into (8), we have (11).

1) With high SNR, the decoder almost provides perfect extrinsic information. Thus, for a single sequence j with $U_1^j = N$, all $\xi_{u_1}^j \to 1$ and other terms are null:

$$p(g) \propto \sum_{i=0}^{q-1} \exp\left\{-\frac{N\mathcal{E}_{\text{av}}}{2\sigma_n^2}|g_i - h|^2\right\} \delta(g - g_i). \quad (12)$$

2) With low SNR, all extrinsic probabilities are close to $1/M$. Thus, in (11), there are M^N items. However, we can consider the following factors:

$$\exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left(1 - X_{12}^{j\,2}\right)|h|^2\right] \triangleq \epsilon_{12} \quad (13)$$

and

$$\exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left(1 - X_{34}^{j\,2}\right)|h|^2\right] \triangleq \epsilon_{34}. \quad (14)$$

With $N \gg 1$, in the items where $X_{12}^{j\,2} + X_{34}^{j\,2} < 2$, $\epsilon_{12}\epsilon_{34} \to 0$. Hence, only sequences with $X_{12}^{j\,2} + X_{34}^{j\,2} \geq 2$ correspond to non-negligable terms in (11). In addition, there are also another two items in (11):

$$\exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left|g_i + X_{12}^j h\right|^2\right] \exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left|g_i + jX_{34}^j h\right|^2\right]. \quad (15)$$

Since the variance $4\sigma_n^2/N\mathcal{E}_{\text{av}}$ of the two Gaussian distributions in (15) are very small, the product can be neglected when they have different mean values. Thus, only the items with $X_{12}^j = 0$ or $X_{34}^j = 0$ are non-negligible in (11). Furthermore, among these non-negligible items, only four items with $U_1^j = N$, $U_2^j = N$, $U_3^j = N$ and $U_4^j = N$ are dominant taking into account of (13) and (14). The item corresponding to $U_u^j = N$ is:

$$\exp\left\{-\frac{N\mathcal{E}_{\text{av}}}{2\sigma_n^2}|g_i - \alpha_u h|^2\right\}, \quad (16)$$

where $u = \{1, 2, 3, 4\}$. So, (11) can be simplified into:

$$p(g) \propto \sum_{i=0}^{q-1}\sum_{u=1}^{4}\left\{\exp\left(-\frac{N\mathcal{E}_{\text{av}}}{2\sigma_n^2}|g_i - \alpha_u h|^2\right) \times \prod_{k=0}^{N-1}\xi_k^{(u)}\right\}, \quad (17)$$

where $\xi_k^{(u)}$ is the extrinsic probability of $\alpha_u x_k$. Especially, when all extrinsic probabilities are close to $1/M$, the four items in (17) have the same factors.

So, combining the approximations with high and low SNR, the pdf of channel estimate with M-QAM in factor graph can be approximated as a mixture of four Gaussian distributions:

$$p(g) \propto \sum_{i=0}^{q-1}\sum_{u=1}^{4}\left\{\beta_u \exp\left(-\frac{N\mathcal{E}_{\text{av}}}{2\sigma_n^2}|g_i - h|^2\right)\right\}, \quad (18)$$

where β_u, $u = 1, \ldots, 4$, represent the products of extrinsic probabilities and $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$. Our simulations confirm this conclusion, as shown in Fig. 4.

IV. APP EVALUATION FROM DOWNWARD MESSAGES

With the conclusions in Section III, the discrete channel distribution can be approximated as a sum of four Gaussian distributions. However, our simulations show that there is always one dominant Gaussian distribution (with mean value h) among the four ones even with low SNR. Hence, when calculating APP, we only consider the dominant one ($\beta_1 = 1$) and the discrete distribution message $\mu_{p(g) \to f_k}$ can be reduced to one pair $(\hat{h}, \hat{\sigma}_h^2)$ characterizing $p(g)$. The parameter $\hat{h}$ is computed as the mean value of the discrete distribution in (2) (resp. (8)) and $\hat{\sigma}_h^2$ is equal to $\sigma_n^2/(L_p\mathcal{E}_{\text{pilot}})$ (resp. $\sigma_n^2/(N\mathcal{E}_{\text{av}})$). Thus, we can calculate each downward message APP_k in a continuous way, instead of computing it for each codebook

$$\sum_{k=0}^{N-1}|s_k^j g_i - h x_k - n_k|^2 \approx \frac{N\mathcal{E}_{\text{av}}}{2}\left[\left|g_i + X_{12}^j h\right|^2 + \left(1 - X_{12}^{j\,2}\right)|h|^2 + \left|g_i + jX_{34}^j h\right|^2 + \left(1 - X_{34}^{j\,2}\right)|h|^2\right] + 2N\sigma_n^2,$$

$$\text{where} \quad X_{12}^j \triangleq \frac{2\left(U_2^j - U_1^j\right)}{N} \quad \text{and} \quad X_{34}^j \triangleq \frac{2\left(U_4^j - U_3^j\right)}{N} \quad (10)$$

$$p(g) \propto \sum_{i=0}^{q-1}\delta(g - g_i)\sum_{j=1}^{M^N}\left\{\exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left|g_i + X_{12}^j h\right|^2\right] \times \exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left(1 - X_{12}^{j\,2}\right)|h|^2\right]\right.$$

$$\times \exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left|g_i + jX_{34}^j h\right|^2\right] \times \exp\left[-\frac{N\mathcal{E}_{\text{av}}}{4\sigma_n^2}\left(1 - X_{34}^{j\,2}\right)|h|^2\right]$$

$$\left.\times \prod_{u_1=0}^{U_1^j-1}\xi_{u_1}^j \prod_{u_2=0}^{U_2^j-1}\xi_{u_2}^j \prod_{u_3=0}^{U_3^j-1}\xi_{u_3}^j \prod_{u_4=0}^{U_4^j-1}\xi_{u_4}^j \prod_{v=0}^{V-1}\xi_v^j\right\}. \quad (11)$$
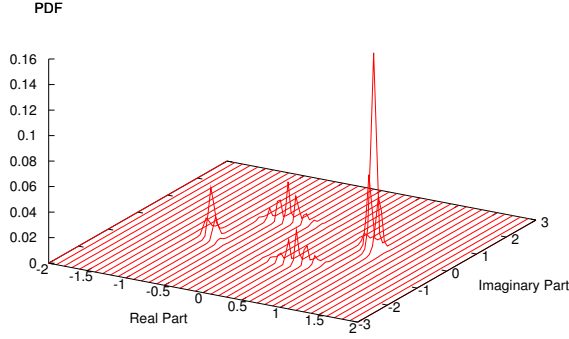
Fig. 4. The pdf of channel estimate distribution based on 16-QAM pilots and data with transmitted $E_s/N_0 = 4$dB. The channel coefficient is 0.9 + j0.4. A half-rate convolutional code (23, 35) is used.
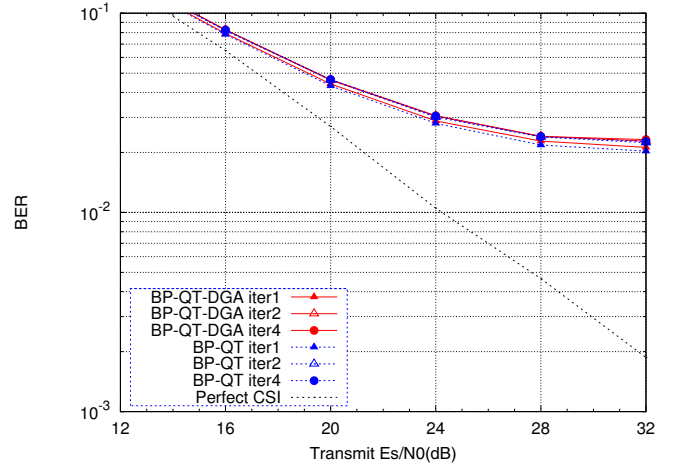

Fig. 5. Bit error rate performance comparison of BP-QT-DGA and BP-QT. Rate 1/2 convolutional code (23, 35) with 16-QAM (2-dimensional Ungerboeck mapping) with block length 260. Number of pilots is 10. The quantization interval is [-3, +3] with a step equal to 0.2.

value g_i, and then marginalizing with respect to g. It reduces the computation complexity.

The $\mathrm{APP}_{k,m}$ that transmitted symbol x_k equals s_m can be expressed as:

$$\mathrm{APP}_{k,m} \propto \int_g p\left(y_k | x_k = s_m, g\right) p\left(g\right) d g. \qquad (19)$$

With some calculation, we have

$$\mathrm{APP}_{k,m} \propto \frac{1}{8\pi\left(\hat{\sigma}_h^2|s_m|^2 + \sigma_n^2\right)} \exp\left\{-\frac{|y_k - s_m\hat{h}|^2}{2\left(\hat{\sigma}_h^2|s_m|^2 + \sigma_n^2\right)}\right\}. \qquad (20)$$

In Fig. 5, BP using APP computation based on the Gaussian approximation (BP-QT-DGA) is compared with BP using APP computation based on the discrete distribution (BP-QT). We observe that BP-QT-DGA performs as well as BP-QT. Nevertheless, thanks to the computation in (20), a single APP computation instead of q computations is performed for each symbol x_k with the Gaussian approximation. Since the derivation of each message $\mu_{f_k \to p(g)}$ still involves q computations, one per codebook value, the global complexity reduction brought by the Gaussian approximation in the downward messages is approximately 50% for large q.

V. ESTIMATION FROM UPWARD MESSAGES

We observe in Fig. 5 that quantization in the upward message leads to an error floor. Thus, we propose to increase the accuracy of $\hat{h}$ using a continuous upward message.

A. Estimation based on messages from pilots

By replacing the quantization codebook $\{g_i\}$ in (2) with a continuous value g and considering normalization, we get the following estimated mean value of g:

$$\hat{h} = \frac{\int_g g\exp\left\{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{L_p-1}|y_{\mathrm{pilot},k} - x_{\mathrm{pilot}}g|^2\right\} dg}{\int_g \exp\left\{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{L_p-1}|y_{\mathrm{pilot},k} - x_{\mathrm{pilot}}g|^2\right\} dg}. \qquad (21)$$

The integration of numerator and denominator in (21) results in the following simple formula:

$$\hat{h} = \frac{x_{\mathrm{pilot}}^*}{L_p|x_{\mathrm{pilot}}|^2}\sum_{k=0}^{L_p-1} y_{\mathrm{pilot},k}. \qquad (22)$$

B. Estimation based on messages from data

Replacing the discrete computation in (6) by an integral, we get the mean value of g:

$$\hat{h} = \frac{\int_g g\prod_{k=0}^{N-1}\left[\sum_{m=0}^{M-1}\exp\left(-\frac{|y_k - s_m g|^2}{2\sigma_n^2}\right)\xi_{k,m}\right]dg}{\int_g \prod_{k=0}^{N-1}\left[\sum_{m=0}^{M-1}\exp\left(-\frac{|y_k - s_m g|^2}{2\sigma_n^2}\right)\xi_{k,m}\right]dg}, \qquad (23)$$

$$\hat{h} = \left\{\sum_{j=1}^{M^N}\Delta_j\frac{2\pi\sigma_n^2\sum_{k=0}^{N-1}s_k^{j*}y_k}{\left(\sum_{k=0}^{N-1}|s_k^j|^2\right)^2}\exp\left(\frac{1}{2\sigma_n^2}\frac{\left|\sum_{k=0}^{N-1}s_k^{j*}y_k\right|^2}{\sum_{k=0}^{N-1}|s_k^j|^2}\right)\right\}\left\{\sum_{j=1}^{M^N}\Delta_j\frac{2\pi\sigma_n^2}{\sum_{k=0}^{N-1}|s_k^j|^2}\exp\left(\frac{1}{2\sigma_n^2}\frac{\left|\sum_{k=0}^{N-1}s_k^{j*}y_k\right|^2}{\sum_{k=0}^{N-1}|s_k^j|^2}\right)\right\}^{-1} \qquad (26)$$

4

in which

$$\prod_{k=0}^{N-1}\left[\sum_{m=0}^{M-1}\exp\left(-\frac{|y_k-s_m g|^2}{2\sigma_n^2}\right)\xi_{k,m}\right]$$
$$=\sum_{j=1}^{M^N}\exp\left(-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1}|y_k-s_k^j g|^2\right)\Delta_j, \qquad (24)$$

where s_k^j represents the value of the kth symbol in sequence j and

$$\Delta_j=\prod_k \xi_k^j. \qquad (25)$$

After some calculations, we obtain (26). In order to reduce the computation in (26), we only consider the item with the largest Δ_j in both numerator and denominator:

$$\hat{h}\approx\frac{\displaystyle\sum_{k=0}^{N-1}s_k^{j_{\max}*}y_k}{\displaystyle\sum_{k=0}^{N-1}|s_k^{j_{\max}}|^2}, \qquad (27)$$

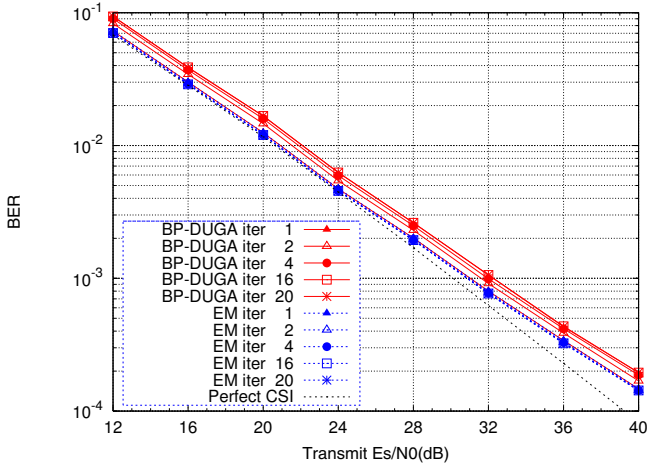where $j_{\max}=\arg\max_j \Delta_j$.

VI. Numerical results

Fig. 6.   Bit error rate performance comparison for channel estimation: BP-DUGA vs EM. Rate 1/2 convolutional code (23, 35) with 16-QAM (Gray mapping) block length 514. Number of pilots is 10 with power 18.

We now compare BP with continuous downward and upward messages (BP-DUGA) with EM. From Fig. 6, we observe that, with 10 pilots, i.e., good initialization, the proposed BP-DUGA does not exhibit an error floor and achieves bit error rate (BER) performance close to EM. Furthermore, from Fig. 7, with 1 pilot for initial estimate, we observe that the proposed BP-DUGA performs better than EM: for BER performance, we obtain a gain of about 1.5 dB with BP-DUGA compared to EM. Using continuous downward and upward messages brings a complexity reduction compared to the quantization method. As a result, BP-DUGA complexity becomes equivalent to EM complexity.
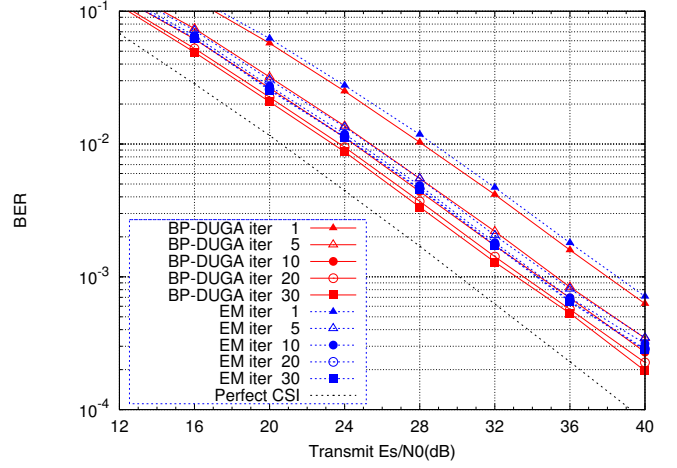

Fig. 7.   Bit error rate performance comparison for channel estimation: BP-DUGA vs EM. Rate 1/2 convolutional code (23, 35) with 16-QAM (Gray mapping) block length 514. Number of pilots is 1 with power 2.

VII. Conclusion

Thanks to an approximation of the distribution of the channel estimate as a mixture of Gaussian distributions, we improved the performance of BP and reduced its complexity by propagating continuous messages in the factor graph. The proposed BP-DUGA almost achieves EM performance under good initialization and outperforms it under a bad initialization. BP-DUGA and EM have equivalent complexity. This paper is focusing on a single-path channel with QAM modulation. Nevertheless, the extension of the Gaussian approximation principle to a multipath channel is natural.

References

[1] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, pp. 28–41, Jan. 2004.

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

[3] H. Wymeersch, H. Steendam, H. Bruneel, and M. Moeneclaey, "Code-aided frame synchronization and phase ambiguity resolution," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2747–2757, Jul. 2006.

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

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society (B)*, vol. 13, no. 6, pp. 1–38, 1977.

[6] A. S. Gallo and G. M. Vitetta, "Soft-in soft-output detection in the presence of parametric uncertainty via the bayesian em algorithm," *EURASIP JWCN*, vol. 2005, no. 2, pp. 100–116, 2005.

[7] F. Z. Merli and G. M. Vitetta, "A factor graph approach to the iterative detection of ofdm signals in the presence of carrier frequency offset and phase noise," in *Proc. IEEE ICC '07*, Jun. 2007, pp. 2865–2870.

[8] J. Dauwels, S. Korl, and H.-A. Loeliger, "Particle methods as message passing," in *Proc. IEEE ISIT '06*, Jul. 2007, pp. 2052–2056.