LDPC Codes for Fading Channels: Two Strategies

Xiaowei Jin, Teng Li, Thomas E. Fuja, and Oliver M. Collins

University of Notre Dame, Notre Dame, IN 46556, USA
{xjin, tli, tfuja, ocollins}@nd.edu

Abstract. This paper compares two approaches to reliable communication over fading channels using low-density parity check (LDPC) codes. The particular model considered is the block fading channel with independent channel realizations. The first approach uses a block interleaver to create independent sub-channels that are encoded using irregular LDPC codes with rates specified by the appropriate capacity values; this first approach uses a decision feedback structure wherein decoded data are used as pilots to estimate the channel prior to the next round of decoding. The second approach uses a combined graph describing both the channel and the code to iteratively estimate the channel and decode data. For comparable channels, it is shown that the decision-feedback approach provides better performance when a long delay is acceptable, while the iterative receiver provides better performance when more stringent delay constraints are imposed.

1 Introduction

Modern error correcting codes such as low-density parity-check (LDPC) codes [1] and turbo codes [2] provide excellent performance over fading channels. In such systems, the receiver must estimate the characteristics of the fading – i.e., the channel state information (CSI) – to effectively decode the data. The optimal approach is to carry out joint channel estimation and decoding; however, the complexity of optimal joint channel estimation and decoding can be prohibitive.

A popular alternative to joint estimation/decoding is to design *iterative receivers* that iteratively estimate the channel and decode the data. An iterative receiver provides a good approximation to the optimal approach with reasonable complexity. A unified approach for designing iterative receivers on factor graphs was proposed by Worthen et. al [5]; this unified approach makes it possible to employ the iterative *sum-product algorithm* on factor graphs [3] describing the iterative receiver. Performance analysis of the resulting receiver is possible by means of density evolution [4]; this also leads to the design of good codes that are well-matched to the receiver. Examples of iterative receivers employing LDPC and turbo codes can be found in [6,7,8]. Although the iterative receiver approach is suboptimal, it achieves good performance.

More recently, a receiver employing decision feedback based successive decoding has been proposed for channels with memory [9]. This approach decomposes

M. Fossorier et al. (Eds.): AAECC 2006, LNCS 3857, pp. 235–244, 2006.

a fading channel (or a channel with memory) into a bank of memoryless sub-channels with a block interleaver, and each sub-channel is protected with an LDPC code. The receiver is composed of a decoder and a channel estimator. The LDPC codes are decoded successively, and the decoded symbols are fed back to the channel estimator, which uses the feedback to estimate the channel and then decode the LDPC code for the next sub-channel. It has been shown that this approach is optimal [9]. However, it can incur a long delay in order to achieve optimal performance.

The goal of this paper is to compare the performance of these two approaches. The channel model considered in this paper is a block fading channel with independent channel realizations between blocks. The transmitted signal is subject to frequency-flat, time-selective fading with both amplitude fluctuation and phase rotation. The complex fading coefficient remains constant for T channel uses and then changes to a new (independent) value for the next T channel uses. (In this paper, we refer to each group of T channel uses over which the channel is constant as a *block*.) Neither the transmitter nor the receiver are assumed to know the channel realization. The effect of a delay constraint is examined for each receiver structure.

2 Channel Model

The receiver produces samples of the matched filter output at the symbol rate. The equivalent discrete-time complex channel model is given by

$$y_{i,k} = c_i x_{i,k} + w_{i,k}, \quad i = 1, 2, \ldots N, \quad k = 1, 2, \ldots, T, \tag{1}$$

where the fading coefficients $\{c_i\}$ are i.i.d. complex Gaussian random variables with distribution $\sim \mathcal{CN}(0, 1)$ and the additive noise $\{w_{i,k}\}$ are similarly i.i.d. complex Gaussian with distribution $\mathcal{CN}(0, N_0)$. Here N_0 is the noise variance per two dimensions. In the equation above, $x_{i,k}$ is the kth transmitted symbol of the ith block and $y_{i,k}$ is the corresponding received symbol. For simplicity, we assume binary phase shift keying (BPSK) modulation, so $x_{i,k} \in \mathcal{S} = \{+1, -1\}$.

3 The Successive Decoding Receiver

To transmit data over a block fading channel with coherence time T, the transmitter de-multiplexes the user data into T streams. Each stream is then individually encoded using a block code of length N and rate R_k for $k = 1, \ldots, T$. The kth codeword is denoted $\overline{\mathbf{x}}_k = [x_{1,k}, \ldots, x_{N,k}]^T$ for $k = 1, \ldots, T$. These codewords are stored column-wise in the following block structure:

$$\begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,k} & \cdots & x_{1,T} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,k} & \cdots & x_{2,T} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i,1} & x_{i,2} & \cdots & x_{i,k} & \cdots & x_{i,T} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,k} & \cdots & x_{N,T} \end{pmatrix}. \tag{2}$$

The transmitter sends the data in 2 row by row. We will also use $\overline{\mathbf{x}}_{i,:}$ to denote the ith row and $\overline{\mathbf{x}}_{:,j}$ to denote the jth columns in (2).

3.1 Estimation and Decoding

The receiver employs a successive decoding strategy that was proposed for channels with memory in [9]. It operates on the block structure in (2), starting from the leftmost column and proceeding to the right. The codeword $\overline{\mathbf{x}}_{:,1}$ is decoded first, and $\overline{\mathbf{x}}_{:,2}$ is decoded second with the assistance of the decoded symbols corresponding to $\overline{\mathbf{x}}_{:,1}$. More specifically, the decoded symbols corresponding to $\overline{\mathbf{x}}_{:,1}$ are used to estimate the channel realizations that affect each symbol in $\overline{\mathbf{x}}_{:,2}$. This approach is used to permit decoding to proceed from left to right.

The estimation and the decoding of a codeword are performed sequentially. Take the kth codeword as an example, where $1 \leq k \leq T$. At this point, all the previous $k-1$ codewords are decoded, and the decoded symbols have been fed back to the receiver. First, the receiver estimates the *a posteriori* probability (APP) of the ith bit in the kth codeword as

$$\mathrm{APP}(x_{i,k} = a) = P\big(x_{i,k} = a | \overline{\mathbf{y}}_{i,:}, x_{i,1}, \cdots, x_{i,k-1}\big) \tag{3}$$

for $a \in \{+1, -1\}$ and $i = 1, \ldots, N$. In (3), the bits $x_{i,1}, \ldots, x_{i,k-1}$ are treated as training symbols. After the receiver calculates the log-likelihood ratios (LLRs) $\{\xi(1,k), \ldots, \xi(N,k)\}$, where

$$\xi(i,k) = \log \frac{\mathrm{APP}(x_{i,k} = +1)}{\mathrm{APP}(x_{i,k} = -1)}, \tag{4}$$

the decoder uses the LLRs to decode the kth LDPC codeword.

3.2 Optimality

The receiver structure described above is optimal, i.e., it is information lossless if the decisions fed back at each stage are correct. This was shown in [9] for any channels with memory. The rest of this section will briefly illustrate the result for the block fading channel.

The main idea is that the block transmission structure effectively decomposes the original block fading channel into a bank of T sub-channels. These sub-channels are memoryless, but they interact with each other via the decision feedback. Thus, the bits in a codeword are transmitted over a memoryless sub-channel, and separate estimation and decoding is optimal. To see this, we write the constrained channel capacity of a block fading channel as

$$C = \frac{1}{TN} I(\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_T; \overline{\mathbf{y}}_1, \ldots, \overline{\mathbf{y}}_T) \tag{5}$$

where N is assumed to be sufficiently large. Now define the kth sub-channel as follows: it has a scalar inputs $x_{i,k}$, a vector output $\overline{\mathbf{y}}_{i,:}$ and a vector of decision feedbacks $[x_{i,1}, \ldots, x_{i,k-1}]^T$. The capacity of the kth sub-channel is given by

$$C_k = \frac{1}{N} I\Big(\overline{\mathbf{x}}_k; \overline{\mathbf{y}}_1, \ldots, \overline{\mathbf{y}}_T \Big| [x_{1,1}, \ldots, x_{1,k-1}]^T, \ldots, [x_{T,1}, \ldots, x_{T,k-1}]^T\Big). \tag{6}$$

From the chain rule of mutual information, we have

$$C = \frac{1}{T} \sum_{k=1}^{T} C_k \tag{7}$$

which means the original channel can be decomposed into T sub-channels without loss of mutual information. Furthermore, due to the independent nature of the original channel, the sub-channel is memoryless, i.e.,

$$C_k = \frac{1}{N} \sum_{i=1}^{N} I(x_{i,k}; \overline{\mathbf{y}}_{i,:} | x_{i,1}, \ldots, x_{i,k-1}) \tag{8}$$

Finally, the APP value in (3) is a sufficient statistics for the sub-channel. Therefore, the estimation and decoding scheme in Section 3.1 is optimal.

Intuitively, the block fading channel with coherence time T and i.i.d. inputs can be viewed as a multiple access channel with T independent users and a vector channel output. Using this analogy, the successive interference cancellation scheme, which is optimal for a multiple access channel, becomes a successive decoding scheme, wherein the decision feedback serve as training symbols.

3.3 The APP Calculation

This section describes how the APP in (3) can be computed. Since the techniques for estimating the APP values are the same for any row of (2), we will only consider the first row and drop the time index i for the rest of the paper. In what follows, x_1^k is used to denote the vector $[x_1, x_2, \ldots, x_k]$, and y_1^k is defined similarly. Since

$$P(x_k = a | y_1^T, x_1^{k-1}) \propto P\big(y_1^T | x_1^{k-1}, x_k = a\big), \tag{9}$$

we will consider the computation of likelihood function (9). Minimum mean square error (MMSE) channel estimation uses the decision feedback to obtain an MMSE estimate of the channel state and then enumerates all possible values of the unknown (or future) symbols to obtain the probability 9. Mathematically,

$$\begin{aligned} P\big(y_1^T | x_1^k\big) &= \sum_{x_{k+1}^T \in \mathcal{S}^{T-k}} P(x_{k+1}^T) P\big(y_1^T | x_1^T\big) \\ &= \sum_{x_{k+1}^T \in \mathcal{S}^{T-k}} P(x_{k+1}^T) P(y_1^{k-1} | x_1^{k-1}) P(y_k^T | y_1^{k-1}, x_1^T) \qquad (10) \\ &\propto \sum_{x_{k+1}^T \in \mathcal{S}^{T-k}} P(y_k^T | y_1^{k-1}, x_1^T) \qquad (11) \\ &= \sum_{x_{k+1}^T \in \mathcal{S}^{T-k}} \frac{1}{|\pi \Sigma|} \exp\Big(- \big(y_k^T - x_k^T \hat{c}\big)^H \Sigma^{-1} \big(y_k^T - x_k^T \hat{c}\big)\Big) \qquad (12) \end{aligned}$$

where from linear estimation theory [10], the conditional mean and variance are given by

$$\hat{c} = \frac{1}{k-1+N_0} \sum_{i=1}^{k-1} y_i x_i^* \quad \text{and} \tag{13}$$

$$\Sigma = \frac{N_0}{k-1+N_0} x_k^T (x_k^T)^H + N_0 \mathbf{I}_{T-k+1}. \tag{14}$$

3.4 Channel Capacity

Using the simplified notation, the constrained capacity of sub-channel k is

$$C_k = I(x_k; y_1^T | x_1^{k-1}). \tag{15}$$

From the definition of mutual information and entropy, (15) becomes

$$C_k = H(x) - \mathrm{E}[-\log \mathrm{APP}(x_k)], \tag{16}$$

where the APP value can be computed using (9) and (12). The expectation in (16) can be evaluated using Monte Carlo integration.

Due to the increasing number of training symbols, the sequence of sub-channel capacity is monotonic increasing, i.e.,

$$C_1 < C_2 < \cdots < C_T. \tag{17}$$

The kth sub-channel is coded by a particular LDPC codes of rate R_k. Here, we set the code rate to be equal to the sub-channel capacity, i.e.,

$$R_k = C_k, \quad \text{for } k = 1, \ldots, T. \tag{18}$$

This paper used irregular LDPC codes optimized for the AWGN channel as component codes.

4 The Iterative Receiver

This section derives an algorithm that carries out iterative channel estimation and LDPC decoding on a joint factor graph. Since the channel is a complex fading channel, pilot symbols are used to assist in estimating the channel states. The basic idea of the iterative receiver is to permit the channel state estimator and the iterative decoder to share preliminary information about the transmitted symbols; after several iterations of LDPC decoding, the decoded symbols are fed back to the channel estimator as additional pilots to help refine the channel estimation.

4.1 Factor Graph

The system factor graph is shown in Figure 1. The LDPC decoder is represented by a bipartite graph in which the variable nodes V represent transmitted symbols and the factor nodes C represent parity checks. One pilot symbol (designated

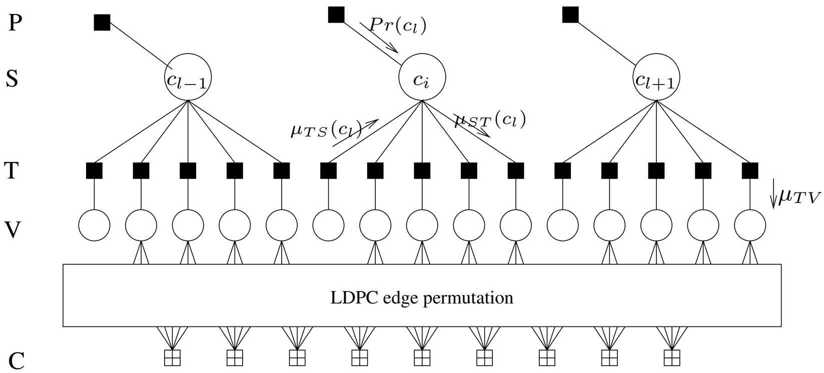

Fig. 1. A fraction of factor graph for a block fading channel with LDPC code, where channel states are complex fading coefficients and independent with each other

x_{l0} for all $l \in \{1, 2, \ldots, N\}$) is transmitted in each fading block; the value of this pilot symbol is known to the receiver and is not part of any LDPC codeword. The joint graph is obtained by concatenating the code graph and the channel graph. In the channel graph, the channel states are denoted by variable nodes S. The factor nodes $T_{lk} = f(y_{lk}|x_{lk}, c_l)$ represent the input-output characteristic of the channel. The factor nodes P represent the prior information about the distribution of the channel state.

In the following, we use notation $\mathcal{CN}(x, m, \sigma^2)$ to represent the complex Gaussian distribution of x with mean m and variance σ^2 per two dimensions. The message from the channel to the decoder and the message from the decoder to the channel are denoted by $\mu_{T_{lk} \to V_{lk}}(y_{lk})$ and $\mu_{V_{lk} \to T_{lk}}(y_{lk})$, respectively; they represent log-likelihood ratios associated with x_{lk}. In contrast, the messages in the channel graph are probability density functions (PDFs). The message from T_{lk} to S_l is $\mu_{T_{lk} \to S_l}(c_l)$, which is the PDF of y_{lk} given the channel state c_l:

$$\mu_{T_{lk} \to S_l}(c_l) = Pr(y_{lk}|c_l, \hat{x}_{lk}) \propto \mathcal{CN}(y_{lk}, c_l\hat{x}_{lk}, \sigma_n^2) \propto \mathcal{CN}(c_l, y_{lk}/\hat{x}_{lk}, \sigma_n^{)}. \quad (19)$$

Here $\hat{x}_{lk}$ represents the decisions made by the channel decoder for the transmitted symbols.

The message from S_l to T_{lk} is the PDF of c_l given the pilots, and is denoted by $\mu_{S_l \to T_{lk}}(c_l)$. Note that, in the first channel estimation iteration, only one pilot symbol is sent for each fading block; however, in the subsequent iterations, additional pilots are obtained by taking hard decisions about the transmitted symbols from the LDPC decoder. Finally the message from P_l to S_l is basically the prior distribution of the channel state c_l, which is $Pr(c_l) = \mathcal{CN}(c_l, 0, 1)$.

4.2 The Iterative Sum-Product Algorithm

Once the messages on the combined factor graph have been defined, it is straightforward to derive the message passing algorithm that iteratively estimates the channel and decodes the LDPC code. Because LDPC decoding via message passing is well understood and widely known, this section will emphasize the aspects of the receiver dealing with channel estimation.

The iterative algorithm works as follows. First, the channel estimator obtains the initial estimate of the channel using the pilot symbols. The LLRs of the channel ($\mu_{T_{lk}\to V_{lk}}(y_{lk})$) are then calculated based on this channel estimate and are provided to the LDPC decoder. After several LDPC decoding iterations, new LLRs ($\mu_{V_{lk}\to T_{lk}}(y_{lk})$) are calculated by the decoder, and hard decisions ($\hat{x}_{lk}, k \neq 0$) of the transmitted symbols are obtained by the channel estimator based on $\mu_{V_{lk}\to T_{lk}}(y_{lk})$. The following hard decision rule is applied,

$$\hat{x}_{lk} = \begin{cases} 1, & \mu_{V_{lk}\to T_{lk}} > T \\ -1, & \mu_{V_{lk}\to T_{lk}} < -T \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

If T is sufficiently large, then the code symbols with messages with absolute values that are greater than T are highly reliable, and the resulting non-zero values of $\hat{x}_{lk}$ can be treated as "pseudo-pilots" to help re-estimate the channel.

According to the sum-product rule, the message produced by a state variable node is the product of the input messages, Thus, the message from the channel state node S_l to the factor node T_{lk} is

$$\begin{aligned} \mu_{S_l\to T_{lk}}(c_l) &= Pr(c_l) \prod_{j=0, j\neq k, \hat{x}_{lj}\neq 0}^{j=N-1} \mu_{T_{lj}\to S_l}(c_l) \\ &= \mathcal{CN}(c_l, 0, 1) \prod_{j=0, j\neq k, \hat{x}_{lj}\neq 0}^{j=N-1} \mathcal{CN}(c_l, y_{lk}/\hat{x}_{lk}, \sigma_n^2) \\ &\propto \mathcal{CN}(c_l, \hat{m}_c, \hat{\sigma}_c^2). \end{aligned} \tag{21}$$

The expressions for $\hat{m}_c$ and $\hat{\sigma}_c^2$ can be obtained by applying the product rule for Gaussian PDFs (see Appendix A of [3]) and are omitted here. Also by the sum-product rule, the message produced by a factor node is the product of the input messages with the local factor, marginalized for the destination variable. Thus the message out of the factor node T_{lk} is

$$\begin{aligned} Pr(y_{lk}|x_{lk}) &= \int_{c_l} \mu_{S_l\to T_{lk}}(c_l) Pr(y_{lk}|c_l, x_{lk}) dc_l \\ &= \int_{c_l} \mathcal{CN}(c_l, \hat{m}_c, \hat{\sigma}_c^2) \mathcal{CN}(y_{lk}, x_{lk}c_l, \sigma_n^2) dc_l \\ &\propto \mathcal{CN}(y_{lk}, x_{lk}\hat{m}_c, \hat{\sigma}_c^2 + \sigma_n^2). \end{aligned} \tag{22}$$

Equation (22) comes from the the integration rule for Gaussian PDFs. (See Appendix of [3]). Since the messages from the channel to the decoder are LLRs, we have

$$\mu_{T_{lk}\to V_{lk}}(y_{lk}) = \log \frac{Pr(y_{lk}|x_{lk}=1)}{Pr(y_{lk}|x_{lk}=-1)} = \frac{2\text{Re}\{y_{lk}^* \hat{m}_c\}}{\hat{\sigma}_c^2 + \sigma_n^2}. \qquad (23)$$

5 Simulation Results and Conclusions

We first considered a block fading channel with block length $T = 5$.In this case, the successive decoding receiver uses five codes of rates $R_1, \ldots, R_5$ set to 0, 0.4948, 0.5643, 0.5917, 0.6058, respectively. The overall rate is 0.4513. For the iterative receiver, a code of rate 0.5641 is used, so taking into account the pilots the overall rate is also 0.4513. The channel capacity in terms of $Eb/N0$ is 5.6 dB. To compare the performance of the two receivers, the overall delay is set to be the same. We set the delay to be 200k, 50k, and 6k bits, respectively. For the successive receiver, the codeword length of each sub-channel is 40k, 10k and 1.2k. For the iterative receiver, the codeword length is 160k, 40k, and 4.8k.

The results of the $T = 5$ block fading channel are plotted in Fig. 2. When the delay is large, the successive decoding scheme outperforms the iterative receiver, while at a short delay, the iterative receiver performs better. Intuitively, when a long delay is acceptable, the successive decoding receiver, proven to be optimal by preserving channel mutual information under the assumption that the fed back decoded symbols are correct, will always performs at least as well as the iterative receiver. On the other hand, since the iterative receiver uses a single code, as compared to T codes used in the successive decoding receiver, the block length of the LDPC code in the iterative receiver is T times that of constituent codes in the successive receiver. (Taking into account the one-bit training symbol for each fading block in the iterative receiver, the exact ratio of the component codeword length is $T - 1$.) When the over all delay is relatively short, this difference in codeword length has significant impact on system performance, as clearly demonstrated in the 6k bits delay curve in Fig. 2. In a moderate delay constraint of 50k bits, the performances of the two approaches are rather close.

We also simulated a $T = 10$ block fading channel. In the simulation, the successive receiver uses 10 codes of rates 0, 0.5177, 0.5869, 0.6109, 0.6229, 0.6302, 0.6364, 0.6397, 0.6430 and 0.6453. The iterative receiver uses a code of rate 0.5014. The overall rate of both system is 0.4513. The performance comparison for delay constraints of 10k, 100k and 200k bits results are shown in Fig. 3. The results are similar to the case of $T = 5$. Note that the performance gap between the iterative and successive receiver increases to around 1 dB in the long delay case. This is due to the fact that if delay is fixed, longer channel memory means less channel diversity, which degrades the performance of the iterative receiver.

In conclusion, the decision-feedback based successive receiver has better capacity approaching performance if long delay is acceptable, while iterative receiver is more robust to delay constraints. Currently we are looking at the comparison of two approaches on more practical channel models. We are also investigating the possibility of combining the two design philosophies into one receiver design, that takes the advantages of both approaches.

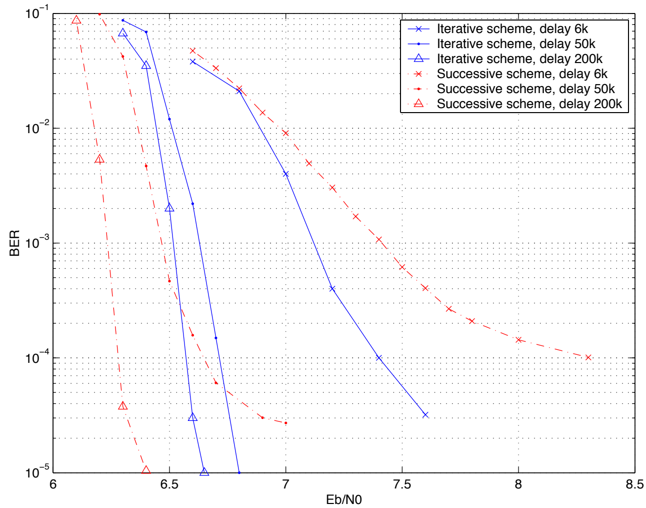

Fig. 2. Performance comparison of successive and iterative schemes for a $T = 5$ independent block fading channel under different delay constraints

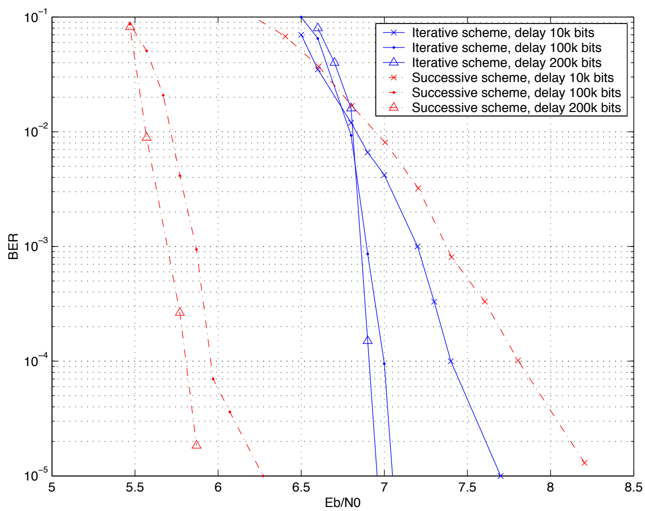

Fig. 3. Performance comparison of successive and iterative schemes for a $T = 10$ independent block fading channel under different delay constraints

References

1. R. G. Gallager: Low Density Parity Check Codes, MIT Press (1963)
2. G. Berrou, A. Gla vieux, P. Thitimajshima: Near shannon limit error-correcting coding: Turbo codes. *Proc. 1993 Int. Conf. Commun.*, pp. 1064-1070, 1993.

3. F. R. Kschischang, B. J. Frey, H.-A. Loeliger: Factor Graphs and the Sum-Product Algorithm. *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
4. T. J. Richardson, R. L. Urbanke: The capacity of Low-Density Parity-Check codes under message-passing decoding. *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618 2001
5. A. P. Worthen, W. E. Stark: Unified design of iterative receivers using factor graphs. *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 843–849, 2001.
6. R.-R. Chen, R. Koetter, U. Madhow, D. Agrawal: Joint demodulation and decoding for the noncoherent block fading channel: a practical framework for approaching Shannon capacity. *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1676-1689, Oct. 2003.
7. K. Fu and A. Anastasopoulos: Analysis and design of LDPC codes for time-selective complex-fading channels. *IEEE Trans. Wireless Communications*, vol. 4, no. 3, pp. 1175-1185, May 2005.
8. X. Jin, A. W. Eckford, T. E. Fuja: Analysis and design of low density parity-check codes for non-coherent block fading channels. *Proc. of the Int. Symp. on Inform. Theory,* Adelaide, Australia, Sep. 2005.
9. T. Li, O. M. Collins: A successive deocding strategy for channels with memory. *Proc. of the Int. Symp. on Inform. Theory*, Adelaide, Australia, Sep. 2005.
10. S. M. Kay: Fundamentals of statistical signal processing estimation theory. Prentice Hall, 1993.