OPEN ACCESS

University of BRISTOL

Link to published version (if available):
10.1109/VETECS.2006.1683066

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

Performance of the EP-MBCJR algorithm in time dispersive MIMO office environments

C. M. Vithanage and C. Andrieu
Department of Mathematics
University of Bristol
Bristol BS8 1TW, UK.
{C.Vithanage,C.Andrieu}@bristol.ac.uk

R. J. Piechocki
Centre for Communications Research
University of Bristol
Bristol BS8 1UB, UK.
R.J.Piechocki@bristol.ac.uk

J. P. Coon
Toshiba Research Europe Limited
32, Queen Square, Bristol BS1 4ND, UK.
J.P.Coon@toshiba-trel.com

Abstract— **This work investigates the performance of a new reduced-complexity trellis decoding algorithm (termed the EP-MBCJR algorithm) when employed for the task of equalization in an indoor multiple input multiple output wireless environment. The algorithm is a generic approximate reduced-state variation of the BCJR algorithm modeled after the conventional M-BCJR algorithm. Instead of choosing the active states based on the filtered distribution of states in the forward recursion, the EP-MBCJR algorithm selects the active states based on "beliefs" on the states. This can be seen as an application of the concept of "Expectation propagation" and leads to identical forward and backward recursions which can be iterated to improve system performance.**

A receiver architecture comprising of channel estimation, sampling phase selection and turbo equalization is proposed and its performance evaluated through computer simulations. For the simulation of channels closely resembling the physical environment, we have used channels generated in accordance with the IEEE 802.11n TGn channel models.

I. INTRODUCTION

In trying to cater for the insatiable demand for data rate, wireless communication systems need to employ capacity enhancing technologies such as multiple input multiple output (MIMO) channels and will need to handle the time dispersion from the intermediate channels which is going to become even more significant with the reduction of the symbol durations. One possible approach for future high data rate communication systems is to consider time domain equalization of MIMO channels. It is well appreciated that optimal and efficient equalization can be performed using the BCJR algorithm of [1], though the resulting system complexity is huge. Various reduced complexity variations of this optimal algorithm have been developed in the recent history, including the M-BCJR algorithm of [2]. The M-BCJR algorithm selects a sequence of active states in the trellis, during the forward recursion and thereby confines the computations on to these active states, reducing the computational complexity. One possible reason for the lack of robustness of the M-BCJR algorithm is that the selection of the active states is based on the filtered distribution of states at each time instant [3].

A family of approximate inference algorithms under the rubric of "Expectation propagation" is presented in [4]. This concept is applied in [5] for inference in general dynamic Bayesian networks. Noting that this algorithm when applied to a hidden Markov model (HMM) enables the use of approximating distributions for the forward and backward messages, in the EP-MBCJR algorithm we apply a variation of [5] for the complexity reduction in the BCJR. More details of this new algorithm can be found in [6]. The new algorithm uses discrete distributions with M significant components only, as the approximating distributions of the forward/backward messages. In short, the application of the new algorithm boils down to selecting the active states of the forward/backward messages in each of the forward/backward iterations based on "beliefs" of the states. Therefore the proposed algorithm is similar to the M-BCJR algorithm with the active state selection criterion being changed from the filtered distribution of states to the beliefs of states. Additionally, there is active state selection in both of the forward and backward recursions of the new algorithm and these can be iterated to achieve improved performance.

After an overview of the simulation model in section II, a receiver structure is presented in section III to implement the tasks of down sampling phase selection, channel estimation and turbo equalization. The proposed reduced complexity trellis decoding scheme is only applied to the equalizer whereas it can be applied for the channel decoders as well. In the computer simulations of the system, we use oversampled channels for the intermediate MIMO channel which conform to the IEEE 802.11n TGn channel model specifications. Simulation results given in section IV show the reduction in computational complexity offered by the EP-MBCJR algorithm. Section V gives the final conclusions.

II. SIMULATION MODEL

We will consider an n_t transmit antenna, n_r receive antenna MIMO communication system in a frequency selective environment with turbo equalization being performed at the receiver, as shown in Fig. 1. Spatial multiplexing or the transmission of independent data streams across the antennas is considered at the transmitter. Assuming the receiver to perform base band sampling of the received signal at a rate of F_s times per symbol instant, we can consider the equivalent simulation model with the transmitted signals and the channel effects oversampled at the same rate [7]. Therefore the simulation model represents the transmitter pulse shaping filters, MIMO
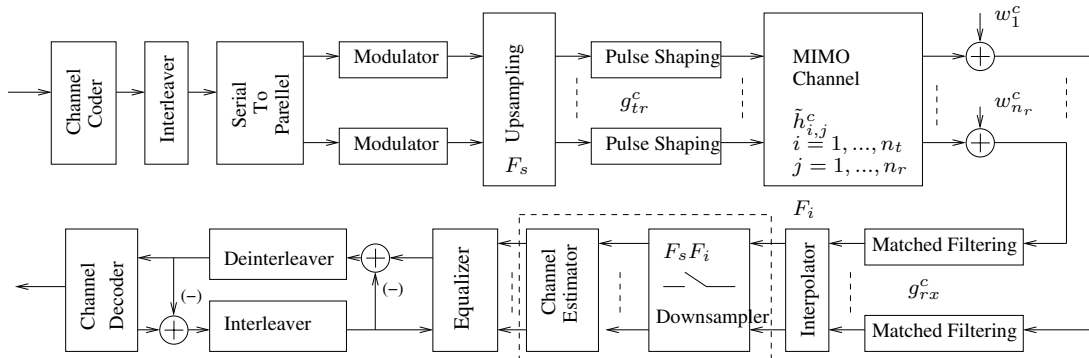
Fig. 1. The Simulation Model

channel impulse responses and the receiver matched filters oversampled to the rate of F_s. The pulse shaping filters at the transmitter are assumed to be square root raised cosine filters.

At the receiver, matched filtering is matched to the transmitter pulse shaping filters and hence has the same square root raised cosine impulse response. After matched filtering, we make provision for interpolation of the sampled signals at the rate of F_i per symbol instant. This resultant signal is downsampled at the rate of $Q = F_s F_i$ to give symbol rate samples at the chosen sampling phase to the remaining signal processing functions of the receiver. In this downsampling operation, the selection of the downsampling phase of the signal (out of Q possible) corresponds to the suitable timing position selection. As shown in the next section, we will select this phase based on the MIMO channels they produce to the receiver. It should be noted that the "channels" as perceived by the receiver will be the impulse responses resulting from the actual intermediate channels as well as the transmit and receive pulse shaping filters. In this paper, we will use an equivalent complex base-band representation for the analysis of signals and use the superscript c to explicitly denote continuous time signals. Denoting the continuous time impulse responses of the transmit ant receive filters as $g_{tr}^c(\tau)$ and $g_{rx}^c(\tau)$ respectively, and the actual intermediate channel between the ith transmit antenna and the jth receive antenna as $\tilde{h}_{i,j}^c(t,\tau)$, the channel perceived by the receiver is

$$h_{i,j}^c(t,\tau) = g_{tr}^c(\tau) * \tilde{h}_{i,j}^c(t,\tau) * g_{rx}^c(\tau).$$

Here, $*$ denotes the convolution operation. Assuming the channel remains invariant during the transmission of a frame of signals, we have $h_{i,j}^c(t,\tau) = h_{i,j}^c(\tau)$.

The equalization and channel decoding tasks operate at the symbol rate and are coupled together to form the standard turbo equalization structure [8].

III. CHANNEL ESTIMATION AND EQUALIZATION

A. Channel estimation and sampling phase selection

When there is no time dispersion from the channel and when the combined transmit-receive filters satisfy the Nyquist criteria for no inter symbol interference (ISI), there exists an

optimal sampling point which produces no ISI [9]. Practically, the sampling point which is closest to this optimal point can be selected, for example, using the method of [10]. As the time dispersion of the intermediate channel increases, the concept of an optimal sampling point disappears and the receiver has to choose the sampling phase based on other criteria which considers subsequent signal processing. Sampling phase selection to minimize the received signal to interference plus noise ratio is suggested in [11], and is more appropriate in the context of decision feedback equalization. In this work, we will select the sampling phase based on the view of the MIMO channels they present to the receiver. Therefore the sampling phase selection and channel estimation stages of the receiver will be combined together.

The continuous time received signal on each receive antenna j (with perfect frequency synchronization) is

$$y_j^c(t) = \sum_{i=1}^{n_t} \sum_{k=-\infty}^{\infty} x_i(k) h_{i,j}^c(t - kT_S) + w_j^c(t). \quad (1)$$

Here, T_S is the symbol duration, k is a discrete symbol epoch index and $x_i(k)$ denotes the complex modulated symbol transmitted by the ith antenna at the kth symbol duration.

We can consider the sampled discrete-time representation of these signals as

$$y_j(k,p) = \sum_{i=1}^{n_t} \sum_{l=0}^{L} x_i(k-l) h_{i,j}^{(p)}(l) + w_j(k,p). \quad (2)$$

We have assumed the channel delay spread to extend to $L + 1$ symbols only and the causality of the channel is justified by considerations of a suitable delay in the receiver and the finite duration of the impulse responses of practical transmit-receive filters. The time indices k and p refer to the actual time instant $(kT_S + (p-1)T_S/Q)$ with $p \in \{1, ..., Q\}$ and $k \in \mathcal{Z}$. ($\mathcal{Z}$ denotes the set of integers.) Therefore p represents the sampling phase which needs to be chosen for the succeeding symbol spaced operations. We can also note that due to the use of root raised cosine filters, $w_j(\bullet, p)$ are independently and identically distributed (i.i.d.) for a given p as $\mathcal{CN}(0, N_0)$ (i.e. circularly symmetric complex Gaussian distributed with a variance of N_0).

We will consider sampling phase selection based on received signal samples pertaining to a training sequence (of T_{train} symbols) which is appended in front of the data sequence in each frame transmission. We can collect the received signals pertaining to a particular sampling phase and the training symbols as

$$\mathbf{Y}^{(p)} = \mathbf{X}\mathbf{H}^{(p)} + \mathbf{W}^{(p)}. \tag{3}$$

Here, $\mathbf{Y}^{(p)} = \left[\widehat{\mathbf{y}}_1(p) \cdots \widehat{\mathbf{y}}_{n_r}(p) \right]$ with $\widehat{\mathbf{y}}_j(p) = [y_j(1,p) \cdots y_j(T_{train},p)]^\dagger$, where $(\bullet)^\dagger$ denotes the matrix transpose operation. Similarly, $\mathbf{W}^{(p)} = \left[\widehat{\mathbf{w}}_1(p) \cdots \widehat{\mathbf{w}}_{n_r}(p) \right]$ with $\widehat{\mathbf{w}}_j(p) = [w_j(1,p) \cdots w_j(T_{train},p)]^\dagger$. $\mathbf{H}^{(p)} = \left[\mathbf{H}^{(p)}(0)^\dagger \cdots \mathbf{H}^{(p)}(L)^\dagger \right]^\dagger$ where $\mathbf{H}^{(p)}(l)$ is the $n_t \times n_r$ matrix with the $(i,j)^{th}$ element being $h_{i,j}^{(p)}(l)$. The kth row of $T_{train} \times n_t(L+1)$ matrix $\mathbf{X}$ is $\left[\widehat{\mathbf{x}}(k) \cdots \widehat{\mathbf{x}}(k-L) \quad \mathbf{0}_{1 \times fill} \right]$, with $\widehat{\mathbf{x}}(k) = [x_1(k) \cdots x_{n_t}(k)]$ and the zero vector $\mathbf{0}_{1 \times fill}$ ensuring the length of each column is $n_t(L+1)$.

Therefore, given the $T_{train}Qn_r$ received signals corresponding to the training symbols only, we can estimate the channels at each sampling phase to minimize the least squares error as

$$\hat{\mathbf{H}}^{(p)} = \left(\mathbf{X}^\ddagger \mathbf{X} \right)^{-1} \mathbf{X}^\ddagger \mathbf{Y}^{(p)} \quad ; \quad p \in \{1, \cdots, Q\}. \tag{4}$$

Here, $(\bullet)^\ddagger$ and $(\bullet)^{-1}$ denote conjugate transpose and inversion operations on a matrix. We assume the above matrix inverse to exist. In addition, the noise variance on each of the receive antennas, corresponding to the pth sampling phase can be estimated as

$$\hat{N}_0^{(p)} = \left\| \mathbf{Y}^{(p)} - \mathbf{X}\hat{\mathbf{H}}^{(p)} \right\|_F^2 \Big/ (T_{train}n_r), \tag{5}$$

where $\|\mathbf{A}\|_F$ denotes the Frobenius norm of matrix $\mathbf{A}$.

We will select the sampling phase based on the maximum of

$$\left\| \hat{\mathbf{H}}^{(p)} \right\|_F^2 \Big/ \hat{N}_0^{(p)}. \tag{6}$$

In simulations, this method presented a bit error rate performance close to a scheme with an exhaustive decoding for all sampling phases and can be viewed as a method of maximizing the received signal to noise energy ratio.

B. Reduced-complexity equalization

Given the selected sampling phase, $\hat{p}$ and the estimated channel state information $\hat{\mathbf{H}}^{(\hat{p})}$ and $\hat{N}_0^{(\hat{p})}$; the received discrete-time symbol spaced signals at each symbol time k and at each receive antenna j can be approximately represented as

$$y_j(k) = \sum_{i=1}^{n_t} \sum_{l=0}^{L} x_i(k-l)\hat{h}_{i,j}(l) + \hat{w}_j(k). \tag{7}$$

From here onwards we use $y_j(k)$ to denote $y_j(k,\hat{p})$ and $\hat{w}_j(k)$ are considered i.i.d. as $\mathcal{CN}(0,\hat{N}_0)$ for all j and k with $\hat{N}_0 = \hat{N}_0^{(\hat{p})}$. Let us denote the frame length by T_{frame} and use the notations $\mathbf{y}(k) = [y_1(k) \cdots y_{n_r}(k)]^\dagger$ and $\mathbf{y} =$

$\left[\mathbf{y}(1)^\dagger \cdots \mathbf{y}(T_{frame})^\dagger \right]^\dagger$. Considering the subsequent channel decoding operation, optimal equalization amounts to the generation of the posterior marginal distributions $p(x_i(k)|\mathbf{y})$ for $k \in \{T_{train}+1, ..., T_{frame}\}$ and $i \in \{1, ..., n_t\}$.

To accomplish this task, it is convenient to define a state for the equalizer at time k consisting of candidates for some transmitted symbols as $\mathbf{s}_k = \left[\widehat{\mathbf{x}}(k) \cdots \widehat{\mathbf{x}}(k-L+1) \right]^\dagger$ and consider the possible evolutions of this state with time, or the *trellis* of the equalizer. With this definition, the BCJR algorithm presented in [1] can be used to efficiently compute the probability mass functions (pmf) $p(\mathbf{s}_{k-1}, \mathbf{s}_k|\mathbf{y})$ for each $k \in \{T_{train}+1, ..., T_{frame}\}$, and from these the required posterior marginal distributions, $p(x_i(k)|\mathbf{y})$ can be computed.

The BCJR algorithm consists of a forward and a backward recursion through the trellis and computes the well known $\alpha(\mathbf{s}_k)$, $\beta(\mathbf{s}_k)$ and $\gamma(\mathbf{s}_{k-1}, \mathbf{s}_k)$ in its two recursions [1]. A final combination of these computations results in the posterior marginal pmf outputs: $p(\mathbf{s}_k|\mathbf{y}) \propto \alpha(\mathbf{s}_k)\beta(\mathbf{s}_k)$ and $p(\mathbf{s}_{k-1}, \mathbf{s}_k|\mathbf{y}) \propto \alpha(\mathbf{s}_k)\gamma(\mathbf{s}_{k-1}, \mathbf{s}_k)\beta(\mathbf{s}_k)$.

The complexity of this optimal algorithm is exponential in both n_t and L since each of the forward and backward messages when normalized are distributions over a set with a configuration space having a size exponential in n_t and L. Hence as a reduced complexity approach we use the reduced complexity trellis decoding method termed the "EP-MBCJR algorithm" proposed by the authors in [6]. This algorithm is a combination of the M-BCJR algorithm [2] and the concept of "Expectation Propagation" as applied in [5], and its improvement over the conventional M-BCJR algorithm is illustrated in [6].

The M-BCJR algorithm of [2] reduces the complexity of the optimal algorithm by using reduced support approximations for the forward and backward messages at each time instant. Denoting the corresponding forward and backward messages of time k as $\tilde{\alpha}(\mathbf{s}_k)$ and $\tilde{\beta}(\mathbf{s}_k)$, each of these distributions have a support (or the number of non-zero probability state configurations) of size M ($M \ll |\mathbf{s}_k|$, with $|\mathbf{s}_k|$ denoting the cardinality of the set of possible configurations of $\mathbf{s}_k$). The support set at time k (denoted Ω_k) is common to both forward and backward messages, and are selected during the forward recursion based on approximations of the forward messages themselves. For example if $\tilde{\alpha}(\mathbf{s}_{k-1})$ and Ω_{k-1} are given it is obvious (and visibly apparent in the trellis structure) that there is a set of states at time k (which we will denote as Ξ_k) that are candidates for becoming the set Ω_k. The M-BCJR algorithm computes temporary forward messages $\hat{\alpha}(\mathbf{s}_k)$ for $\mathbf{s}_k \in \Xi_k$ and simply chooses as Ω_k the set of M states with the largest $\hat{\alpha}(\mathbf{s}_k)$.

The EP algorithm of [5] is applicable to a more general dynamic Bayesian network. When applied to a Markov chain type network as in the case of the HMM, this algorithm leads to identical forward and backward recursions which can be iterated. Also the algorithm makes provision for the use of approximating distributions for the messages that will lead to a reduction in complexity of their representations,

as long as these approximations are in some exponential family of distributions. If $\tilde{\alpha}^u(\mathbf{s}_{k-1})$ and $\tilde{\beta}^{u-1}(\mathbf{s}_k)$ are such approximating distributions at the relevant iterations u and $u-1$, the computation of $\tilde{\alpha}^u(\mathbf{s}_k)$ is as follows:

1) Compute $\hat{\alpha}(\mathbf{s}_k) = \sum_{\mathbf{s}_k} \tilde{\alpha}^u(\mathbf{s}_{k-1}) \gamma(\mathbf{s}_{k-1}, \mathbf{s}_k)$.
2) Compute $\hat{p}(\mathbf{s}_k | \mathbf{y}) \propto \hat{\alpha}(\mathbf{s}_k) \tilde{\beta}^{u-1}(\mathbf{s}_k)$.
3) Project $\hat{p}(\mathbf{s}_k | \mathbf{y})$ onto $q(\mathbf{s}_k)$ which is in some exponential family of distributions by minimizing the Kullback-Leibler divergence $D(\hat{p} \| q) = \sum_{\mathbf{s}_k} \hat{p}(\mathbf{s}_k | \mathbf{y}) \{\log(\hat{p}(\mathbf{s}_k | \mathbf{y})) - \log(q(\mathbf{s}_k))\}$.
4) Finally develop the new forward messages as $\tilde{\alpha}^u(\mathbf{s}_k) = q(\mathbf{s}_k) / \tilde{\beta}^{u-1}(\mathbf{s}_k)$.

Again, the fact that the developed approximations of the forward/backward messages are in some exponential family of distributions leads to the reduction in complexity compared to an optimal algorithm.

In the EP-MBCJR algorithm [6] we make two main modifications to the ideas of [5]. First we relax the requirement of $q(\mathbf{s}_k)$ being in some exponential family of distributions and instead use the family of discrete distributions with a support restricted to M components. Secondly instead of minimizing the divergence $D(\hat{p} \| q)$ in the projection step, we minimize the divergence $D(q \| \hat{p})$.

As described in [6], this leads to the generic reduced-complexity trellis decoding algorithm as described in Table I below. Effects of the known training sequence are neglected for brevity. Only the forward recursion in the uth iteration is described since all the recursions are identical in nature. The candidates for the support and the selected support of time k in the uth iteration are denoted $\Xi^u_{k,fwd/bwd}$ and $\Omega^u_{k,fwd/bwd}$ respectively with a further indication of the direction of recursion. A terminated trellis (at state $\mathbf{s}_0$) is assumed. Also we have used the parameter η as a positive minimum set to the forward/backward messages. This enables the possibility of forward/backward messages with different supports and the proposed algorithm was found to have a bit error rate performance which is invariant for a wide range of η [6].

IV. SIMULATION RESULTS

We have performed the computer simulations on a $n_t = n_r = 2$ system operating at a carrier frequency of F_c. The root raised cosine filters had a roll off factor of 0.3 and they were simulated with a finite span of 6 symbol durations. The MIMO channels used in the simulations were generated using the Matlab implementation of the IEEE 802.11n TGn channel models (which uses the experimentally verified Kronecker product model for the specification of the transmit receive spatial correlations [12]) done by L. Schumacher in association with AAU-Csys, FUND-INFO and the IST-2000-30148 I-METRA project. The antennas at each terminal were considered to be half a wavelength apart. Only the short term fading effects were considered without considering any path loss or shadowing effects. Also the Doppler effects were not considered and the transmitter and receiver were considered

TABLE I
u^{th} FORWARD RECURSION OF THE EP-MBCJR ALGORITHM

$\mathcal{A}$:	Set $k=1$, $\Omega^u_{0,fwd} = \{\mathbf{s}_0\}$, $\tilde{\alpha}^u(\mathbf{s}_0 = \mathbf{s}_0) = 1$ and $\tilde{\alpha}^u(\mathbf{s}_0 \neq \mathbf{s}_0) = 0$.
$\mathcal{B}$:	Compute $\hat{\alpha}^u(\mathbf{s}_k) = \sum_{\mathbf{s}_{k-1} \in \Omega^u_{k-1,fwd}} \tilde{\alpha}^u(\mathbf{s}_{k-1}) \gamma(\mathbf{s}_{k-1}, \mathbf{s}_k)$ for $\mathbf{s}_k \in \Xi^u_{k,fwd}$.
$\mathcal{C}$:	Compute $\tilde{p}(\mathbf{s}_k | \mathbf{y}) \propto \hat{\alpha}^u(\mathbf{s}_k) \tilde{\beta}^{u-1}(\mathbf{s}_k)$ for $\mathbf{s}_k \in \Xi^u_{k,fwd}$ with the assumption $\tilde{\beta}^{u-1}(\mathbf{s}_k) = \eta$ for $\mathbf{s}_k \in \Xi^u_{k,fwd} \backslash \Omega^{u-1}_{k,bwd}$.
$\mathcal{D}$:	Select the M largest components of $\tilde{p}(\mathbf{s}_k | \mathbf{y})$ and thereby determine the set $\Omega^u_{k,fwd}$.
$\mathcal{E}$:	Update the forward messages $\tilde{\alpha}^u(\mathbf{s}_k) = \hat{\alpha}^u(\mathbf{s}_k)$ for $\mathbf{s}_k \in \Omega^u_{k,fwd}$
$\mathcal{F}$:	If $k < T_{frame}$, increase k by one and go to step $\mathcal{B}$; otherwise end forward recursion.

to be in a none line of sight (NLOS) situation. Uniformly distributed random sequences were employed as the training sequences.
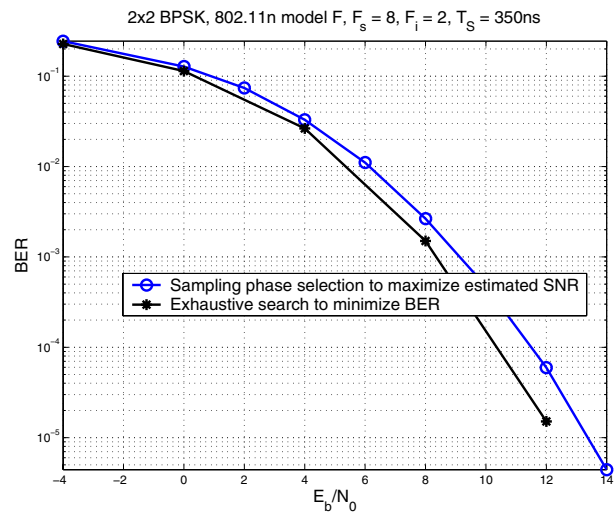

Fig. 2. Comparison of the proposed sampling phase selection with an exhaustive scheme to minimize the decoded bit error rate (BER).

Figure 2 investigates the performance of the sampling phase selection scheme used. No channel coding was employed and a 2×2 BPSK system was investigated when simulated in the IEEE 802.11n TGn channel model F. Also, $F_c = 2GHz$, $F_s = 8, F_i = 2, T_S = 350ns, L = 4, T_{train} = 64$ symbols and $T_{frame} = 384 + L$ symbols. It should be emphasized that it is only the *receiver* which is modeling the channel to have a memory of 4 symbols. For the equalization algorithm, the near optimal algorithm MBCJR-PDA-1 of [3] was utilized with $M = 16$. The proposed sampling phase selection scheme is compared against an exhaustive and impractical scheme which performs decodings for each possible sampling phase and

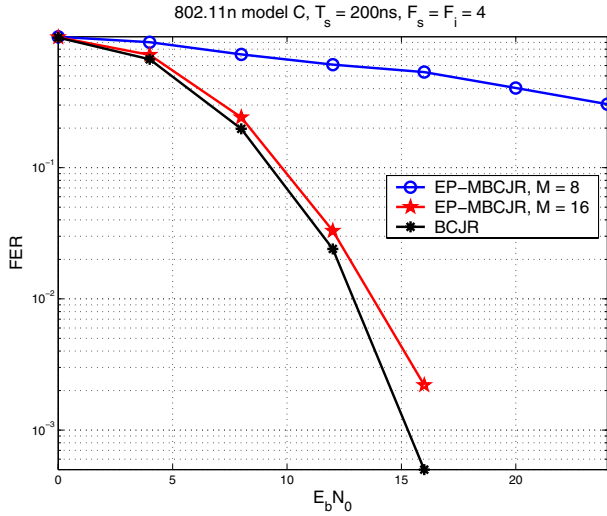selects the decoding which results in the lowest BER. Figures

Fig. 3. Performance of EP-MBCJR scheme in the small office environment. Optimal equalizer has 256 states in the trellis per time instant.
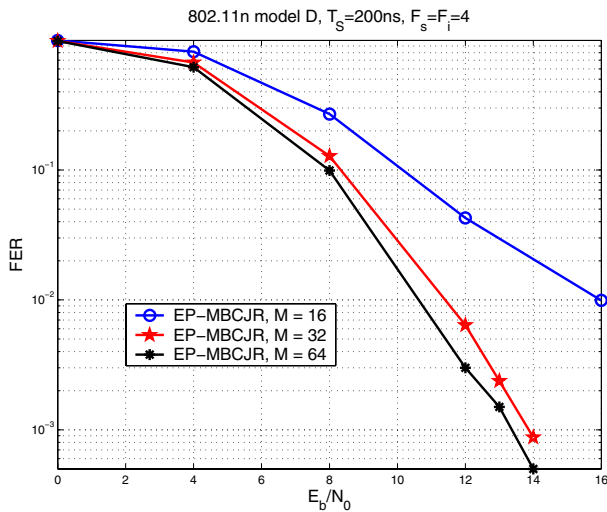

Fig. 4. Performance of EP-MBCJR scheme in the typical office environment. Optimal equalizer has 4096 states in the trellis per time instant.

3 and 4 show the simulations in IEEE 802.11n TGn channel models C and D respectively, with the receiver employing the EP-MBCJR algorithm in a turbo equalization structure. These refer to small office and typical office scenarios. These two scenarios were assumed to have $L = 2$ and $L = 3$ symbol durations, respectively. QPSK modulation was used on each transmit antenna. Other simulation parameters were: $F_c = 5.25GHz$, $F_s = F_i = 4$, $T_S = 200ns$, $T_{train} = 32$ symbols and $T_{frame} = 96 + L$ symbols. A rate half turbo code with two constituent $(5, 7)_8$ convolutional codes was used as the channel code. The receiver performed 5 global turbo equalization iterations. The equalizer performed two sets of forward/backward recursions of the EP-MBCJR algorithm in the first global iteration and performed one set

of forward/backward recursions afterward. During the second global iteration and onwards, the equalizer in its forward recursion used the backward recursion of the previous global iteration. The channel decoder performed 4 turbo decoding iterations per global iteration.

Simulation results illustrate that the EP-MBCJR algorithm is able to achieve a bit error rate performance close to the optimal BCJR algorithm at a much reduced system complexity.

V. CONCLUSIONS

This work implements the generic reduced complexity trellis decoding algorithm termed the EP-MBCJR algorithm for the task of equalization of a MIMO system, which is implemented in indoor environments. A suitable receiver structure involving sampling phase selection, channel estimation and turbo equalization is proposed. The EP-MBCJR algorithm is seen as a preferable alternative to BCJR, even under these more practical scenarios.

REFERENCES

[1] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.

[2] V. Franz and J. B. Anderson, "Concatenated decoding with a reduced-search BCJR algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 186–195, February 1998.

[3] C. M. Vithanage, C. Andrieu, R. J. Piechocki, and M. S. Yee, "Reduced complexity equalization of MIMO systems with a fixed-lag smoothed M-BCJR algorithm," in *Proc. Sixth IEEE Workshop in Signal Processing Advances in Wireless Communications*, New York, USA, June 2005, pp. 155–159.

[4] T. P. Minka, *A family of algorithms for approximate Bayesian inference*, Ph.D. thesis, Massachusetts Institute of Technology, January 2001.

[5] T. Heskes and O. Zoeter, "Expectation propagation for approximate inference in dynamic bayesian networks," in *Proceedings of the Conference on uncertainity in artificial intelligence*, A. Darwiche and N. Friedman, Eds., 2002, pp. 216–233.

[6] C. M. Vithanage, C. Andrieu, and R. J. Piechocki, "Approximate inference in hidden Markov models using iterative active state selection," *IEEE Signal Processing Letters*, vol. 13, no. 2, February 2006.

[7] C. Xiao, J. Wu, S. Y. Leong, Y. R. Zheng, and K. B. Letaief, "A discrete-time model for triply selective MIMO Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1678–1688, September 2004.

[8] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *European Transactions on Telecommunications*, vol. 6, pp. 507–511, September-October 1995.

[9] J. G. Proakis, *Digital Communications*, McGraw–Hill Inc., Singapore, fourth edition, 2001.

[10] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1459–1478, October 1998.

[11] X. Zhu and R. D. Murch, "Layered Space-Time equalization for wireless MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 1189–1203, November 2003.

[12] J. P. Kermoal, L. Schumacher, K. I. Pederson, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1211–1226, August 2002.