Turbo Coding Performance in OFDM Packet Transmission

By: R. Hoshyar, S. H. Jamali, and A. R. Bahai

Dept. of Elect. & Computer Eng., University of Tehran

Email: hoshyar@sanapro.com and jamali@ece.ut.ac.ir

Abstract: *In transmission of the packets using OFDM (Orthogonal Frequency Division Multiplexing) system recommended by IEEE.802.11 [1] there is a constraint on the interleaving depth due to the maximum allowable packet size and processing delay requirement. This results in a non-ideal interleaving which effectively limits the maximum achievable diversity from the channel. In such a channel, fading occurs in block wise manner, i.e., all transmitted symbols of the same block sense the same fade. Also there is considerable correlation among the fade blocks. We call such a channel as Correlated Block Fading (CBF) channel. With the excellent performance of turbo code over AWGN, and ideally interleaved fading channels in mind, we tried to get similarly good performance by using turbo code in OFDM packet transmission. We observed that the system severely suffers from the block fading behavior of the channel. A performance analysis is presented for CBF systems considering the practical case of channel estimation error. Simulations for the OFDM system using rate 1/2 convolutional code and turbo code are performed. While the turbo code has not any gain over the convolutional code with K=7 in BER (Bit Error Rate) performance, it brings about 1.3 dB gain in PER (Packet Error Rate) performance which is more important in packet transmission.*

Index: OFDM, Block Fading, Turbo Coding, BICM, and Channel Estimation Error

I. INTRODUCTION

In communication over a fading channel, it is necessary to use an appropriate interleaving to mitigate and disperse the bursty destructive effect of the channel. For large coherence time or equivalently low Doppler spread of the fading, high interleaving depth is required to break the memory of the channel, which is not affordable in some communication scenarios. In packet switched communications such as the OFDM (Orthogonal Frequency Division Multiplexing) system recommended by IEEE.802.11 [1], or in other applications like GPRS [2] there is a constraint on the interleaving depth due to the maximum allowable packet size or processing delay requirement. In such cases the channel can be modeled as a Block Fading (BF) channel [3], [4], and [5]. In BF model, transmitted sequence is divided into blocks and all the symbols belonging to the same block sense the same fade. In some cases, the fade blocks are assumed to be independent from each other, but in some applications such as the OFDM system there is a considerable correlation among the fade blocks. In this paper we consider the Correlated Block Fading (CBF) channel and present a performance analysis for the coded systems employing bit interleaving between encoder and mapper with an arbitrary signal constellation over the CBF channel.

In section II the model of a coded system using CBF channel, CBF System (CBFS), is described briefly. In section III we consider the OFDM system, and show that it can be modeled by CBFS. Section IV is devoted to the performance analysis of CBFS, and presents an upperbound for the pairwise error probability. Simulation results are presented in section V. In this section we present the performance of turbo code in OFDM packet transmission. Finally, in section VI we make the necessary conclusions.

II. CBFS MODEL

Consider the model for the coded system over the CBF channel illustrated in Figure (1). The k source bits $\underline{u}=(u_1,u_2,\cdots,u_k)$, are encoded into a codeword $\underline{c}=(c_1,c_2,\cdots,c_n)$. Then the coded bits are interleaved and mapped to the L symbol vectors $\{\underline{x}_l\}_{l=1}^{L}$, $\underline{x}_l=(x_{l1},x_{l2},\cdots,x_{lM})$, where x_{li} is a symbol of the signal constellation S, which is used for modulation. Each symbol vector $\underline{x}_l$ is transmitted through a block fading subchannel with a multiplicative distortion α_l, which is a complex Gaussian random variable and is constant during the transmission over the corresponding subchannel. The fading coefficients α_l are correlated with correlation matrix $\mathbf{C}_\alpha = E\left[\underline{\alpha}\underline{\alpha}^h\right]$, $\underline{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_L)^t$.

The received output vectors $\{\underline{y}_l\}_{l=1}^{L}$ are given by:

$$\underline{y}_l=\alpha_l\underline{x}_l+\underline{n}_l, \qquad l=1,\cdots,L, \qquad (1)$$

where the noise vector $\underline{n}_l$ has zero mean independent Gaussian components with variance σ_n^2. We assume that an estimation of the Channel State Information (CSI), i.e. $\underline{\alpha}$, is performed at the receiver. Using the estimated channel state, the received vectors are de-mapped and then de-interleaved according to the specified interleaving pattern. Metrics of the coded bits are computed and passed to an ML (Maximum Likelihood) decoder. Due to the use of bit interleaving the metric computation is suboptimum for non-binary modulations and the decoder is only ML in the sense of working on the given metrics of the coded bits.

III. OFDM SYSTEM MODEL

Figure (2) shows FFT (Fast Fourier Transform) implementation of the OFDM system [6]. Using the notation of Figure (1), the coded vectors $\{\underline{x}_l\}_{l=1}^{L}$ are generated by appropriate coding, interleaving, and mapping. Using IFFT (Inverse Fast Fourier Transform) each vector $\underline{x}_l$ modulates the corresponding subcarrier, such that time domain samples $s_{k,t}$ are as follows:

$$s_{k,t} = G_T \sum_{l=1}^{L} x_{l,t}\, exp\left(j2\pi \frac{kp_l}{N_{FFT}} \right), \qquad k = 0,\cdots,N_{FFT}-1, \qquad t = 1,\cdots,M, \tag{2}$$

where G_T is the IFFT gain factor, p_l determines the corresponding subcarrier of the *l*-th vector $\underline{x}_l$, N_{FFT} is the number of FFT points, and $s_{k,t}$ is the *k*-th sample of the *t*-th OFDM symbol. To prevent ISI, a cyclically extended guard interval (GI), where each OFDM symbol is preceded by a periodic extension of the signal itself, is added. To have an acceptable level of adjacent channel interference, appropriate wave shaping is also applied. The transmit filter h_T is also used to reduce the Adjacent Channel Interference (ACI) by filtering the wideband OFDM signal. The resulted signal is transmitted over the multipath fading cannel with profile (g_p, τ_p), $p = 1,\cdots,P$, where g_p, and τ_p are the gain and delay of the p-th path with $\sum_{p=1}^{P} g_p^2 = 1$, and P is the number of paths. After filtering and GI removing at the receiver side, the received samples $r_{k,t}$ are obtained as follows:

$$\begin{aligned} r_{k,t} &= G_T \sum_{l=1}^{L}\sum_{p=1}^{P} g_p \beta_p x_{l,t} H\left(j2\pi \frac{p_l}{N_{FFT}} \right) exp\left(j2\pi \frac{(k-\tau_p)p_l}{N_{FFT}} \right) + \tilde{v}_{k,t}, \\ &= G_T \sum_{l=1}^{L} x_{l,t}\alpha_l\, exp\left(j2\pi \frac{kp_l}{N_{FFT}} \right) + \tilde{v}_{k,t}, \end{aligned}$$

$$k = 0,\cdots,N_{FFT}-1, \qquad t = 1,\cdots,M. \tag{3}$$

where β_p's are the complex zero mean Gaussian independent random variables with unit variance $E[\beta_p \beta_p^*] = 1$. Due to the small packet size compared to the coherence time of fading (maximum Doppler frequency of 10 Hz), the fade condition of the multipath channel does not change significantly during the transmission of a packet and β_p's are assumed to be constant.

$H(\cdot)$ is the frequency response of the cascade of the transmit and receive filters, and $\tilde{v}_{k,t}$ are the filtered version of noise samples $v_{k,t}$, filtered by receive filter h_R. $v_{k,t}$'s are independent zero mean Gaussian noise components with variance σ_v^2. Finally α_l's are the frequency domain fading coefficients defined as follows:

$$\alpha_l = H\left(j2\pi \frac{p_l}{N_{FFT}}\right)\cdot\left(\sum_{p=1}^{P} g_p \beta_p \exp\left(-j2\pi \frac{\tau_p p_l}{N_{FFT}}\right)\right),\ l=1,\cdots,L. \tag{4}$$

Applying the FFT transform to the received samples $r_{k,t}$, we obtain:

$$y_{l,t} = G_R \sum_{k=0}^{N_{FFT}-1}\left[G_T \sum_{i=1}^{L} \alpha_i x_{i,t}\, exp\left(j2\pi \frac{kp_i}{N_{FFT}}\right) + \tilde{v}_{k,t}\right] exp\left(-j2\pi \frac{kp_l}{N_{FFT}}\right),$$

$$= \alpha_l x_{l,t} + n_{l,t}, \qquad l=1,\cdots,L, \qquad t=1,\cdots,M. \tag{5}$$

where G_R is the FFT gain factor. To normalize the energy of the first part of the received symbol $y_{l,t}$ to 1, we set $G_T = 1/\sqrt{L}$, and $G_R = \sqrt{L}/N_{FFT}$. Assuming that the receive filter h_R has little effect on the inband noise $v_{k,t}$ such that $\tilde{v}_{k,t} \approx v_{k,t}$, the frequency domain noise samples $n_{l,t}$ are related to the time domain noise samples $v_{k,t}$ as follows:

$$n_{l,t} = G_R \sum_{k=0}^{N_{FFT}-1} v_{k,t}\, exp\left(-j2\pi \frac{kp_l}{N_{FFT}}\right), \qquad l=1,\cdots,L, \qquad t=1,\cdots,M, \tag{6}$$

and the variance σ_v^2 is related to the given E_b/N_0:

$$\sigma_v^2 = \frac{N_t}{2N_{DBPOS}\, E_b/N_0}. \tag{7}$$

In the above equation N_{DBPOS} is the number of information bits transmitted with each OFDM symbol, and $N_t = N_{FFT} + N_{GI}$ is the total number of time samples in an OFDM symbol, where N_{GI} is the number of GI samples. It is easy to show that the frequency domain noise samples are also uncorrelated and their variance $\sigma_n^2 = 1/2\, E\left[n_{l,t} n_{l,t}^*\right]$ is as follows:

$$\sigma_n^2 = \frac{1}{2R_t\ E_b/N_0}. \qquad (8)$$

R_t is the total coding rate which in terms of the binary coding rate R_c, number of coded bits per modulation symbol N_{CBPMS}, GI overhead percent $R_{GI} = N_{FFT}/N_t$, and pilot overhead percent $R_{pilot} = L/(L + L_{pilot})$, with L_{pilot} as the number of pilots used in the system, is expressed as follows:

$$R_t = R_c N_{CBPMS}\ R_{GI} R_{pilot}. \qquad (9)$$

Considering equation (6) we can model the OFDM system as a system with CBF channel. The whole blocks, from IFFT to FFT in Figure (2), can be gathered as the CBF channel of Figure (1). In the OFDM system the fade vector $\underline{\alpha}$, and its covariance matrix $\mathbf{C}_\alpha = E\left[\underline{\alpha}\underline{\alpha}^h\right]$ is expressed as follows:

$$\underline{\alpha} = \mathbf{A}\underline{\beta}, \qquad \mathbf{A} = \left[a_{l,p}\right]_{L\times P}, \text{ with } \quad a_{l,p} = g_p\ exp\left(-j2\pi\frac{\tau_p p_l}{N_{FFT}}\right)H\left(j2\pi\frac{p_l}{N_{FFT}}\right),$$

$$\mathbf{C}_\alpha = \mathbf{A}\mathbf{A}^h. \qquad (10)$$

Therefore for the performance analysis of the OFDM system we can use the general model of Figure (1). In the sequel we present a method for the performance evaluation of a bit interleaved coded system over the CBF channel.

IV. Bit Interleaved Coded Modulation

In communication over channels with memory, interleaving has an important role by matching the channel behavior to the code structure. An interleaver with an appropriate size and pattern can completely break the channel memory. For systems using non-binary modulations, it is shown that the appropriate place for the interleaver is between the encoder and mapping units [7], and [8], performing bit interleaving over the coded bits. Due to the use of bit interleaving such systems are called as Bit Interleaved Coded Modulation (BICM) systems. Using such a technique in radio channels, gives better performance, through the increase in system diversity and distance product parameters, compared to the traditional Coded Modulation (CM) systems. This technique will also be useful in OFDM systems, where each subchannel acts as a fading channel, and by interleaving of the coded bits, system tries to distribute the bits in both dimension of time and frequency to make the best use of the possible diversity of the channel. Use of BICM is standardized in IEEE802.11 WLAN for OFDM packet transmission.

The conventional method for decoding of BICM system is composed of sub optimum bit metrics computation, deinterleaving, and then a conventional decoder which is ML in the sense of input bit metrics only. For the conventional receivers Gray mapping has been found appropriate [7], and [8]. However, use of interleaver between the encoder and mapping units introduces memory to the system, such that for ML decoding the both constraints of mapping and coding must be optimized, which is of impractical complexity. Fortunately, bringing the concept of iterative decoding of turbo codes, system can approach to an ML performance using an iterative method [9], and [10].

In scenarios that the channel memory is very large, for example for low Doppler spread of very high bit rate radio links, or there is a limitation on maximum packet size, interleaver is unable to completely break the channel memory, and channel acts as a block fading one. In such channels use of bit interleaving may be debatable, and in some cases traditional CM may perform better. However as the analysis of BICM systems is more general and can include the CM systems analysis as a part. We consider bit interleaving of coded bits in our analysis of systems over CBF channels.

V. PERFORMANCE ANALYSIS OF THE CBFS

While the performance analysis of BICM systems with conventional detection is presented thoroughly in [7], and [8], but they have not considered the practically important effect of channel estimation error. Here we extend the analysis of BICM systems to consider this effect. Then we will us its results in CBFS performance analysis.

V.A. PAIRWISE ERROR UPPERBOUND FOR BICM SYSTEMS

For the performance analysis we use Chernoff bounding technique and compute an upper bound for the pair-wise error probability. We assume Rx antenna diversity with R Rx antennas and channel estimation error. As shown in [10] performances of iterative and conventional decodings are the same if Gray labeling is used in mapper. Then we use the theoretic model presented in this reference (Figure 3). In this model effect of bit interleaving and iterative decoding is seen as random modulation. Here $S_j(\underline{c})$ is subset of mapping constellation S with two elements; one with label $\underline{c}$ and the other with a label resulting from complementing j-th position of the label $\underline{c}$. We denote by $x_j^b(\underline{c})$ that element of $S_j(\underline{c})$ that its j-th label bit is equal to b. The binary mapping $S_j(\underline{c})$ is selected by random selection of j and $\underline{c}$. The ML decoder prefers the codeword $\underline{\hat{C}}$ to the sent codeword $\underline{C}$ if its accumulated metric is greater than that of $\underline{C}$. Therefore, Similar to [7] the pair-wise error probability of choosing the codeword $\underline{\hat{C}}$ instead of the sent codeword $\underline{C}$, conditioned on the channel state information (fading channel gains) can be written as:

$$P\left(\underline{C}\to\underline{\hat{C}}\middle|\underline{P}\right)=\min_{\lambda} E_{\underline{X}|\underline{C}}\, E_{\underline{Y}|\underline{X},\underline{P}}\left\{\exp\left[\lambda\sum_{i=1}^{N}\sum_{r=1}^{R}\frac{1}{\sigma_{n,r}^2}\left(\left|y_{r,i}-\hat{\rho}_{r,i}x_i\right|^2-\left|y_{r,i}-\hat{\rho}_{r,i}\hat{x}_i\right|^2\right)\right]\right\},$$

$$\underline{C}=\{c_1,c_2,\cdots,c_N\}, \tag{11}$$

where $\underline{X}$, $\underline{Y}$, and $\underline{P}$ are the transmitted, received, and channel gain sequences, respectively. λ is the Chernoff parameter. $y_{r,i}=\rho_{r,i}x_i+n_{r,i}$ is the symbol received at time i by Rx antenna r, and $\rho_{r,i}$, and $n_{r,i}$ are the corresponding channel gain, and additive noise, respectively. We assume the general case of different Rx antenna SNRs with $n_{r,i}$ being a zero mean i.i.d.

complex Gaussian process with $Var(\mathrm{Re}(n_{r,i})) = Var(\mathrm{Im}(n_{r,i})) = \sigma_{n,r}^2$. Because of the random selection of the signal constellation, a specified coded sequence $\underline{C}$ may be mapped to different transmitted sequences $\underline{X}$. Therefore, for pair-wise error probability, the expectation must be taken over all possible transmitted sequences $\underline{X}$ ($E_{\underline{X}|\underline{C}}$), and also over all possible noise sequences ($E_{\underline{Y}|\underline{X},\underline{P}}$). Conditioned on channel state information, channel will be memoryless and we have:

$$P(\underline{C} \to \hat{\underline{C}} | \underline{P}) = \min_{\lambda} \prod_{i=1}^{N} E_{x_i|c_i} \prod_{r=1}^{R} E_{y_{r,i}|x_i,\rho_{r,i}} \left\{ \exp\left[\frac{\lambda}{\sigma_{n,r}^2} \left(|y_{r,i} - \hat{\rho}_{r,i} x_i|^2 - |y_{r,i} - \hat{\rho}_{r,i} \hat{x}_i|^2 \right) \right] \right\},$$

$$= \min_{\lambda} \prod_{i=1}^{N} \frac{1}{m 2^m} \sum_{j=1}^{m} \sum_{\underline{c}=1}^{2^m} \prod_{r=1}^{R} E_{y_{r,i}|x_i = x_j^{c_i}(\underline{c}),\rho_{r,i}} \left\{ \exp\left[2\frac{\lambda}{\sigma_{n,r}^2} \mathrm{Re}\left(y_{r,i}^* \hat{\rho}_{r,i} (\hat{x}_i - x_i) \right) + \frac{\lambda}{\sigma_{n,r}^2} |\hat{\rho}_{r,i}|^2 \left(|x_i|^2 - |\hat{x}_i|^2 \right) \right] \right\}$$

,

(12)

where the second line shows averaging over all possible signal constellations $S_j(\underline{c})$. For convenience of notation we have used $x_i = x_j^{c_i}(\underline{c})$, and $\hat{x}_i = x_j^{\hat{c}_i}(\underline{c})$. We first try to perform inner expectation over the additive noise and channel estimation error. We also assume that the channel estimation error $\varepsilon_{r,i} = \hat{\rho}_{r,i} - \rho_{r,i}$ is an i.i.d. zero mean complex Gaussian process with variances $Var(\mathrm{Re}(\varepsilon_{r,i})) = Var(\mathrm{Im}(\varepsilon_{r,i})) = \sigma_{\varepsilon,r}^2$, and independent of the noise and random modulation processes. This assumption is justified for all linear data aided estimation techniques, where estimation error has a linear relation with additive Gaussian noise. Also this is reasonable through the central limit theorem for the estimations consisting of some type of averaging or low pass filtering. For performing the expectation we use the following proposition:

PROPOSITION 1: Assume $X = X_I - jX_Q$ is a complex circular Gaussian noise with mean m_X and variances $Var(X_I) = Var(X_Q) = \sigma^2$, then for real λ_1 and complex λ_2 we have:

$$E\left\{\exp\left[\lambda_1|X|^2+\mathrm{Re}(\lambda_2 X)\right]\right\}=\frac{1}{1-2\lambda_1\sigma_X^2}\exp\left[\frac{\lambda_1|m_X|^2+\mathrm{Re}(\lambda_2 m_X)+|\lambda_2|^2\sigma_X^2/2}{1-2\lambda_1\sigma_X^2}\right]. \tag{13}$$

Considering that the statistics of the random processes are independent of their time index *i*, and also for simplicity of notation we drop the subscript *i*. Using the above proposition for taking expectation over additive noise we reach to the following expression:

$$E_{\varepsilon_r}E_{n_r}\left\{\exp\left[2\frac{\lambda}{\sigma_{n,r}^2}\mathrm{Re}\left(y_r^*\hat{\rho}_r(\hat{x}-x)\right)+\frac{\lambda}{\sigma_{n,r}^2}|\hat{\rho}_r|^2\left(|x|^2-|\hat{x}|^2\right)\right]\right\}=$$

$$E_{\varepsilon_r}\left\{\exp\left[2\frac{\lambda}{\sigma_{n,r}^2}\mathrm{Re}\left(\rho_r^*x^*\hat{\rho}_r(\hat{x}-x)\right)+\frac{\lambda}{\sigma_{n,r}^2}|\hat{\rho}_r|^2\left(|x|^2-|\hat{x}|^2\right)+2\frac{\lambda^2}{\sigma_{n,r}^2}|\hat{\rho}_r|^2|\hat{x}-x|^2\right]\right\}. \tag{14}$$

Using the proposition again for expectation of channel estimation error, and after some manipulations we have:

$$E_{\varepsilon_r}E_{n_r}\left\{\exp\left[2\frac{\lambda}{\sigma_{n,r}^2}\mathrm{Re}\left(y_r^*\hat{\rho}_r(\hat{x}-x)\right)+\frac{\lambda}{\sigma_{n,r}^2}|\hat{\rho}_r|^2\left(|x|^2-|\hat{x}|^2\right)\right]\right\}=$$

$$\frac{1}{1-2\zeta_{\varepsilon,r}\left(\lambda^2W_1+\lambda W_2\right)}\exp\left\{\frac{2\lambda^2\left(1+\zeta_{\varepsilon,r}|x|^2/E_s\right)-\lambda}{1-2\zeta_{\varepsilon,r}\left(\lambda^2W_1+\lambda W_2\right)}\cdot\frac{|\rho_r|^2|\hat{x}-x|^2}{\sigma_{n,r}^2}\right\}, \tag{15}$$

where $E_s=E|x|^2$ is the transmitted signal energy, and also we have used $W_1\overset{\Delta}{=}2|\hat{x}-x|^2/E_s$, $W_2\overset{\Delta}{=}\left(|x|^2-|\hat{x}|^2\right)/E_s$, and $\zeta_{\varepsilon,r}\overset{\Delta}{=}E_s\sigma_{\varepsilon,r}^2/\sigma_{n,r}^2$. Now above expression must be minimized by choosing an appropriate value for λ. This value must be independent of *r*, ρ_r, x, and $\hat{x}$, and also should result in a simple form for pairwise error probability. For acceptable system performance $\sigma_{n,r}^2$, and $\sigma_{\varepsilon,r}^2$ should be small. Assuming low values for these parameters we see that the variation of the above expression in terms of λ mostly depends on the numerator of the exponential function. Setting the derivative of it with respect to λ to zero we get the value

$\lambda = 1/4\left(1+\zeta_{\varepsilon,r}|x|^2/E_s\right)$. This value depends on x and r. Assuming that the first antenna has the largest SNR and using the average of $|x|^2$, we choose $\lambda_{opt} = 1/4\left(1+\zeta_{\varepsilon,1}\right)$. Substituting this value in above equation, we get:

$$E_{\varepsilon_r} E_{n_r}\left\{\exp\left[2\frac{\lambda}{\sigma_{n,r}^2}\mathrm{Re}\left(y_r^*\hat{\rho}_r(\hat{x}-x)\right)+\frac{\lambda}{\sigma_{n,r}^2}|\hat{\rho}_r|^2\left(|x|^2-|\hat{x}|^2\right)\right]\right\}\Bigg|_{\lambda=\lambda_{opt}} =$$

$$\frac{1}{1-2\zeta_{\varepsilon,r}W_{x,\hat{x}}}\exp\left\{-\frac{\mu_{x,r}}{1-2\zeta_{\varepsilon,r}W_{x,\hat{x}}}\cdot\frac{|\rho_r|^2|\hat{x}-x|^2}{8\sigma_{n,r}^2\left(1+\zeta_{\varepsilon,1}\right)}\right\}, \tag{16}$$

where

$$W_{x,\hat{x}} = \frac{|\hat{x}-x|^2}{8E_s\left(1+\zeta_{\varepsilon,1}\right)^2}+\frac{|x|^2-|\hat{x}|^2}{4E_s\left(1+\zeta_{\varepsilon,1}\right)},$$

$$\mu_{x,r} = 2-\frac{1+\zeta_{\varepsilon,r}|x|^2/E_s}{1+\zeta_{\varepsilon,1}}. \tag{17}$$

Now substituting above results in equation (12), we get:

$$P\left(\underline{C}\to\hat{\underline{C}}\middle|\underline{P}\right) = \prod_{i=1}^{N}\left[D\left(\rho_{1,i},\ldots,\rho_{R,i}\right)\right]^{c_i\oplus\hat{c}_i}, \tag{18}$$

where:

$$c_i\oplus\hat{c}_i = \begin{cases}0 & ,c_i=\hat{c}_i\,,\\ 1 & ,c_i\neq\hat{c}_i\,.\end{cases}$$

$$D\left(\rho_{1,i},\ldots,\rho_{R,i}\right)=\frac{1}{m2^m}\sum_{j=1}^{m}\sum_{\underline{c}=1}^{2^m}\prod_{r=1}^{R}\frac{1}{1-2\zeta_{\varepsilon,r}W_{x,\hat{x}}}\exp\left\{-\frac{\mu_{x,r}}{1-2\zeta_{\varepsilon,r}W_{x,\hat{x}}}\frac{|\rho_{r,i}|^2|\hat{x}-x|^2}{8\sigma_{n,r}^2\left(1+\zeta_{\varepsilon,1}\right)}\right\}. \tag{19}$$

For computation of unconditioned pairwise error probability, the expectation over the channel gain must be applied to upper bound (18). This expectation can be readily performed for memoryless Rician and Rayleigh fading channels. For Rayleigh fading channel each $\rho_{r,i}$ is i.i.d.

zero mean circular complex Gaussian process with $Var(\mathrm{Re}(\rho_{r,i})) = Var(\mathrm{Im}(\rho_{r,i})) = 1/2$. Applying expectation to (18) and using the above proposition, we get:

$$P(\underline{C} \rightarrow \hat{\underline{C}}) = \overline{D}^{R \cdot d_H(\underline{C}, \hat{\underline{C}})},$$

$$\overline{D} = \left[\frac{1}{m 2^m} \sum_{j=1}^{m} \sum_{\underline{c}=1}^{2^m} \prod_{r=1}^{R} \frac{1}{1 - 2\zeta_{\varepsilon,r} W_{x,\hat{x}} + \mu_{x,r} \left|\hat{x} - x\right|^2 / \left(8\sigma_{n,r}^2 \left(1 + \zeta_{\varepsilon,1}\right)\right)} \right]^{1/R}. \quad (20)$$

Upper bound for bit error rate (BER) can be computed by using the union bound technique [11]. For convolutional codes it is easily computed using the code generating function. Figure (4) shows the BER simulation and upper bound for a BICM-8PSK coded system over memoryless Rayleigh fading channel. The two cases of ideal channel state information (CSI) and estimation of CSI with $\zeta_\varepsilon = 0.5$ is considered. A terminated rate ½ convolutional code with octal code generator (7,5) is used. A randomly generated pattern of size 2400 is used for bit interleaving. As it is seen from the figure a close agreement exists between upper bounds and simulation results. While expectation of channel gains is readily done for memoryless channels, it is rather complex to be used in the form of (19) for correlated block fading channels. To make it more simple we can use following approximation [Appendix A]:

$$D(\rho_{1,i}, \ldots, \rho_{R,i}) \approx \eta^R \exp\left\{- \sum_{r=1}^{R} \gamma_r \left|\rho_{r,i}\right|^2 \right\}, \quad (21)$$

where γ_r is interpreted as a modified SNR value for Rx antenna *r*, which has the effects of channel estimation errors and random modulation nature of BICM inside. It should be noted that this parameter has a nonlinear dependence to true SNR.

V.B. Pairwise error upper bound for correlated block fading channels

For correlated block fading channels all the symbols transmitted over the same sub blocks will have the same channel gain, to go further on computation of (18), first all of the terms corresponding to the same sub block must be factored and averaged over the same channel gain.

It should be mentioned that also that in such cases channel estimation is performed per sub block and therefore channel estimation error is also same for all of them. This implies that the expectation performed in (16) must be deferred until the factorization of the sub block. Unfortunately it makes the analysis unmanageable. Indeed by performing expectation over channel estimation error before factorization we assume better performance on channel estimation and the results will be to some extent optimistic. We continue our analysis with (21), and express the pairwise error probability conditioned on sub blocks fading as follows:

$$P\left(\underline{C} \to \underline{\hat{C}} \middle| \{\underline{\alpha}_r\}_{r=1}^{R}\right) \approx \eta^{R\, d_H\left(\underline{C},\underline{\hat{C}}\right)} \prod_{r=1}^{R} \prod_{i=1}^{N} \left[\exp\left\{-\gamma_r \left|\alpha_{l(i)}\right|^2\right\}\right]^{c_i \oplus \hat{c}_i} , \tag{22}$$

where $\oplus$ denotes the addition in GF(2), and $l(i)$ is the subchannel over which the i-th coded bit is transmitted by being mapped to a symbol $x_{l,t}$. Similar to [12], the above equation can also be expressed in following matrix form:

$$P\left(\underline{C} \to \underline{\hat{C}} \middle| \{\underline{\alpha}_r\}_{r=1}^{R}\right) \approx \eta^{R\, d_H\left(\underline{C},\underline{\hat{C}}\right)} \prod_{r=1}^{R} \exp\left(-\gamma_r \underline{\alpha}_r^h \mathbf{W} \underline{\alpha}_r\right). \tag{23}$$

In above equation $\underline{\alpha}_r$ is the sub blocks (or sub channels) fading vector for Rx antenna r, $d_H\left(\underline{C},\underline{\hat{C}}\right)$ is the Hamming distance of the two corresponding codes, and $\mathbf{W}$ is the diagonal matrix expressing distribution of Hamming distance of them over the sub blocks:

$$\mathbf{W} = diag\left(\underline{w}\right), \quad \underline{w} = \left(w_1, w_2, \cdots, w_L\right), \quad w_l = \sum_{i=1, l(i)=l}^{n} \left(c_i \oplus \hat{c}_i\right). \tag{24}$$

The fade vector $\underline{\alpha}$ is Gaussian distributed with correlation matrix given in (10). By averaging over $\underline{\alpha}$, and after some manipulation, we can express the unconditional pairwise error probability as follows:

$$P\left(\underline{C} \to \underline{\hat{C}}\right) \approx \eta^{R \cdot d_H\left(\underline{C},\underline{\hat{C}}\right)} \Big/ \prod_{r=1}^{R} \det\left(\mathbf{I}_P + \gamma_r \mathbf{W}_\beta\right), \tag{25}$$

where $\mathbf{I}_P$ is the $P \times P$ identity matrix, and $\mathbf{W}_\beta = \mathbf{A}^h \mathbf{W} \mathbf{A}$. As observed effect of channel estimation error is a decrease in effective SNR through parameter γ_r and also introducing Hamming distance measure through parameter η. This final effect depends really on the value of η, and for values near 1 this effect is negligible. The above equation is also written as follows:

$$P\left(\underline{C} \rightarrow \underline{\hat{C}}\right) = \eta^{R \cdot d_H\left(\underline{C}, \underline{\hat{C}}\right)} \prod_{r=1}^{R} \prod_{p=1}^{P} \frac{1}{1 + \gamma_r \lambda_{p,\beta}(\underline{w})}, \tag{26}$$

where $\lambda_{p,\beta}(\underline{w})$, $p = 1, \cdots, P$ are the eigenvalues of the matrix $\mathbf{W}_\beta$, and depend on the distance distribution vector $\underline{w}$ through the matrix $\mathbf{W}$, and on the channel profile, h_T, and h_R filters through the matrix $\mathbf{A}$. Due to the Hermitian symmetry of the matrix $\mathbf{W}_\beta$, all of its eigenvalues $\lambda_{p,\beta}(\underline{w})$ are positive real values. Equation (26) is very informative; it shows that maximum achievable diversity from each Rx antenna is limited by the number of paths *P*. For the specified profile of the channel this diversity is achieved only when the coding system produces good weight distribution vectors, and SNR is so high that the resulted $\gamma_r \lambda_{p,\beta}(\underline{w})$ are made appropriately large. Due to the linearity of the code we can assume that the transmitted sequence is the all zero codeword, and consider the weight profile of the code instead of its distance profile. For a given E_s/N_0, and weight distribution vector $\underline{w}$, the effective diversity of the codeword $\underline{C}$, can be defined as follows:

$$L_{eff,r}(\underline{C}; E_s/N_0) = \text{Number of } \lambda_{p,\beta}(\underline{w})\text{'s with } \gamma_r \lambda_{p,\beta}(\underline{w}) >> 1. \tag{27}$$

We can define the effective diversity of the code as the minimum of the effective diversity of the nonzero codewords:

$$L_{eff}(E_s/N_0) = \min_{\underline{C} \neq \underline{0}} \sum_{r=1}^{R} L_{eff,r}(\underline{C}; E_s/N_0). \tag{28}$$

The maximum attainable diversity is determined by the corresponding eigenvalues $\lambda_{p,\beta}(\underline{w})$. These values depend both on the channel characteristics, i.e., profile of the channel, and weight distribution.

Figure (5) shows the pairwise error probability $P(\underline{0} \rightarrow \underline{C})$ simulation and upperbound for the minimum weight codeword (11011111001011) of a convolutional code with rate 1/2, and generator polynomial G=(554, 744). We have considered the system model of Figure (2) for the OFDM system with BPSK modulation specified in IEEE.802.11 Draft with channel, and filter parameters given in tables (1) and (2). Interleaving is rectangular and performed over an OFDM symbol with an interleaver depth of 16. The minimum Hamming weight for the specified code is 10. Four cases for possible combinations of RX antenna diversity (equal SNRs), and channel estimation error are considered. Two OFDM training symbols are used for channel estimation resulting channel estimation factor $\zeta_{\varepsilon,1} = \zeta_{\varepsilon,2} = 0.5$ [Appendix B]. As it is seen there is close agreement between simulation and upperbound, which verifies the computed upperbound. Also figure (6) shows the similar results for 16QAM modulation. Here the efficiency of approximation of (21) is observable. While this approximation is not an upperbound, but it is more useful than the loose upper bound resulted by placing the worst exponential instead of all in (19), i.e. $E_i b_i \, e^{-a_i x} \leq (E_i b_i) e^{-\min(a_i)x}$. These two curves coincide for OFDM-BPSK in figure (5).

To compute the BER (Bit Error Rate) upper bound, we can use the union bound:

$$Pb \leq \frac{1}{2}\frac{1}{k} \sum_{\underline{C} \neq \underline{0}} i(\underline{C}) P_2(\underline{C}), \qquad (29)$$

where k is the number of uncoded bits, $i(\underline{C})$ is the number of nonzero uncoded bits in codeword $\underline{C}$, $P_2(\underline{C}) = P(\underline{0} \rightarrow \underline{C})$ is the pairwise error probability, and the multiplicative factor 1/2 is added to tighten the bound [13]. Because of the weight distribution dependency of the pairwise error probability, computation of the union bound through the generating function technique is not possible. We tried to compute the BER upperbound approximately by enumerating the low weight codewords up to a weight limit and truncating the union bound of (29). Unfortunately the results are not satisfactory. Figure (7) shows the BER simulation and truncated union bound for the two rate 1/2 convolutional codes with constraint lengths K=7, and K=9 and corresponding octal generators (554,744), and (561,753), respectively. For K=7 the upperbound is computed with truncation weights 14, 16, and 18, while there is only slight difference between these upperbounds, there is a mismatch between them and simulation. This mismatch is more obvious for K=9. This shows that there are high weight codewords with badly distributed patterns, which deteriorate the system performance.

VI. SIMULATION RESULTS

We performed several simulations for various constraint lengths of the convolutional codes for the OFDM-QPSK system according to the specifications of tables (1) and (2), and using the general model of Figure (1). The results show that both the multipath channel characteristics, and also the code structure are the determining factors in the performance of the system. In the cases that code is good enough, the channel characteristics will be the limiting factor, and conversely there are cases that the code, due to its weak structure, limits the performance of the system. Figure (8) shows the BER simulation results of OFDM-QPSK with packet length of 12 OFDM symbols for different convolutional codes with constraint lengths of K=3, 5, 7, and 9 and octal code generators of (7,5), (46,72), (554,744), and (561,753), respectively. The results show that increasing constraint length K up to 7 improves the BER performance. But there is no significant improvement by increasing K from 7 to 9. This shows that the limiting factor in performance for the values of K greater than 7 is the channel behavior, not the code structure. But for the values of K lower than 7 this is converse and the code structure is the limiting factor.

We used also turbo code in OFDM-QPSK system. The applied turbo code consists of two rate 1/2 recursive systematic convolutional (RSC) codes with octal code generator (1,5/7). To have a rate 1/2 turbo code, the second code systematic bit is punctured completely and the parity bits of the two codes are punctured in turn. Figure (9) shows the BER and PER simulation of OFDM-QPSK with packet length of 64 OFDM symbol for the specified turbo code and rate 1/2 convolutional code with K=7 and octal code generator of (554,744). Interleaver pattern of turbo code is selected randomly. In iterative decoding of turbo code the Max-Sum version of the MAP algorithm (SubMAP) [14] is used up to 10 iterations. Simulation of figure (9) is carried out for two cases: perfect knowledge of $\underline{\alpha}$ (ideal CSI), and estimation of $\underline{\alpha}$ using the two OFDM training symbols according to the IEEE 802.11 standard. As it is seen there is considerable performance loss between ideal CSI and estimation of CSI. This may stir the mind to improve the estimation procedure of CSI, for example with increasing the number of training symbols.

While these two codes have nearly the same BER performance, performance of turbo code at PER 10^{-3} has about 2.2 dB, and 1.3 dB gain for the two cases of ideal CSI, and estimated CSI, respectively. The important point to note is the more sensitivity of turbo code to the CSI estimation error. This may be due to the iterative decoding which bases its work on the symbols reliability values and errors on these values will incur considerable performance loss. However, in the both cases of ideal CSI and estimated CSI, turbo code has better PER performance. Therefore, considering that PER is the main determining factor in the performance of OFDM packet transmission system, it is justified to use turbo code in this system.

VII. CONCLUSION

This paper considered the coded systems using CBF channels. The OFDM system specified by IEEE 802.11 was briefly described, and we showed that OFDM is also in the category of CBFS. We presented the performance analysis of CBFS, and computed an upper bound for the pairwise error probability, in the general case of Rx antenna diversity and error in channel estimation. It is shown that channel estimation has SNR degradation effect. Simulation results showed that the performance of the system is limited by the code structure and channel profile. The maximum achievable diversity is determined by the channel profile, which is achieved by good enough codes. Due to the lack of diversity resource, in OFDM packet transmission frequency domain is the only diversity resource, performance improvement by strengthening the code is hardly achieved. However, In the powerful turbo code compared to the convolutional code with K=7, while has no improvement in BER, presents about 2 dB gain in PER performance which is more important in packet transmission. Therefore, it is justified to use turbo code in OFDM Packet transmission. It is expected also that by using diversity increasing methods such as antenna diversity, turbo code brings about more considerable gain.

VIII. APPENDICES

APPENDIX A: APPROXIMATING SUM OF EXPONENTIALS WITH ONE EXPONENTIAL.

The parameter D in pairwise error probability upperbound (19) is sum of exponentials. While it is manageable in its present form for ideal interleaving of fading channels, it has inconvenient form to be used easily in block fading channels where the assumption of ideal interleaving is not valid further. For this case the approximation of D with a single exponential will be very useful. Here we consider this issue by finding approximation $b\prod_r e^{-a_r x_r}$ for $E_i \prod_r b_{r,i}\, e^{-a_{r,i} x_r}$, where the operator E_i stands for average over index i. We assume all parameters are positive reals. To make such an approximation, we should adopt a distribution for x_r's. This depends on the problem to be applied. The exponential distribution $f_X(x_1,\ldots,x_R)=\prod_r \beta_r\, e^{-\beta_r x_r}$ seems to be useful for its simple form and also adjusting of its parameters β_r for appropriate emphasis on important values of x_r's. Note that for β_r=1, corresponds to the distribution of $\left|\rho_{r,i}\right|^2$ in our analysis. We consider mean squared error as a measure and choose parameters a_r and b to minimize it:

$$
\begin{aligned}
E &= E_X \left| b \prod_r e^{-a_r x_r} - E_i \prod_r b_{r,i}\, e^{-a_{r,i} x_r} \right|^2, \\
&= E_X \left[b^2 \prod_r e^{-2a_r x_r} - 2b\, E_i \prod_r b_{r,i}\, e^{-\left(a_r + a_{r,i}\right) x_r} + E_i E_j \prod_r b_{r,i} b_{r,j}\, e^{-\left(a_{r,i} + a_{r,j}\right) x_r} \right], \\
&= b^2 \prod_r \frac{\beta_r}{2a_r + \beta_r} - 2b\, E_i \prod_r \frac{b_{r,i}\beta_r}{a_r + a_{r,i} + \beta_r} + E_i E_j \prod_r \frac{b_{r,i} b_{r,j} \beta_r}{a_{r,i} + a_{r,j} + \beta_r}. \qquad \text{(A1)}
\end{aligned}
$$

Differentiating this measure with respect to parameters a_r and b, and setting them equal to zero we have:

$$
\begin{aligned}
b &= E_i \left(\prod_r b_{r,i} \frac{2a_r + \beta_r}{a_r + a_{r,i} + \beta_r} \right), \\
b &= E_i \left(\frac{2a_t + \beta_t}{a_t + a_{t,i} + \beta_t} \prod_r b_{r,i} \frac{2a_r + \beta_r}{a_r + a_{r,i} + \beta_r} \right), \qquad t = 1, \ldots, R. \qquad \text{(A2)}
\end{aligned}
$$

The above two equations can be solved numerically. Now we can approximate D in (19) by choosing:

$$
b_{r,i} = \frac{1}{1 - 2\zeta_{\varepsilon,r} W_{x,\hat{x}}}, \qquad a_{r,i} = \frac{\mu_{x,r}}{1 - 2\zeta_{\varepsilon,r} W_{x,\hat{x}}} \frac{\left|\hat{x} - x\right|^2}{8\sigma_{n,r}^2 \left(1 + \zeta_{\varepsilon,1}\right)}, \qquad \text{(A3)}
$$

where i stands for the two indices of j and c. With such an approximation D will take the following form:

$$
D\left(\rho_{1,i}, \ldots, \rho_{R,i}\right) \approx \eta^R \exp\left\{ -\sum_{r=1}^{R} \gamma_r \left|\rho_{r,i}\right|^2 \right\}, \qquad \text{(A4)}
$$

where γ_r can be interpreted as an SNR value that the effects of channel estimation errors and random modulation nature of BICM have been considered in it.

APPENDIX B: ESTIMATION OF THE OFDM SUBCHANNEL GAINS.

Suppose that M_α OFDM training symbol is devoted for estimation of fade vector $\underline{\alpha}$. We denote the M_α received training symbol from *l*-th subchannel by $\underline{y}_l = \left(y_{l1}, y_{l2}, \cdots, y_{lM_\alpha}\right)^t$. According to equation (5) we can write:

$$\underline{y}_l = \alpha_l \underline{x}_l + \underline{n}_l , \qquad l = 1, \cdots, L , \tag{B1}$$

where $\underline{x}_l = \left(x_{l1}, x_{l2}, \cdots, x_{lM_\alpha}\right)^t$ is the transmitted training symbols through the *l*-th subchannel, and $\underline{n}_l = \left(n_{l1}, n_{l2}, \cdots, n_{lM_\alpha}\right)^t$ is the frequency domain Gaussian zero mean noise vector in the *l*-th subchannel. It has uncorrelated components with variance σ_n^2. Using the Gaussian distribution of noise vector, the logarithm of the likelihood function $\Lambda_l\left(\underline{y}_l ; \alpha_l\right) = p\left(\underline{y}_l \middle| \underline{x}_l, \alpha_l\right)$ can be expressed as follows:

$$ln\ \Lambda_l\left(\underline{y}_l ; \alpha_l\right) \propto -\left(\underline{y}_l - \alpha_l \underline{x}_l\right)^h \left(\underline{y}_l - \alpha_l \underline{x}_l\right). \tag{B2}$$

The ML estimate $\hat{\alpha}_l$ is the value that maximizes the above value. By setting $\dfrac{\partial\, ln\, \Lambda_l\left(\underline{y}_l ; \alpha_l\right)}{\partial \alpha_l}$ equal to zero (α_l is a complex value) the estimate $\hat{\alpha}_l$ is computed as follows:

$$\hat{\alpha}_l = \frac{\underline{x}_l^h \underline{y}_l}{\underline{x}_l^h \underline{x}_l} . \tag{B3}$$

The estimation error $\delta_l = \alpha_l - \hat{\alpha}_l$ is given by:

$$\delta_l = -\frac{\underline{x}_l^h \underline{n}_l}{\underline{x}_l^h \underline{x}_l} . \tag{B4}$$

Due to the whiteness, and Gaussian distribution of the noise vector $\underline{n}_l$, δ_l's, $l = 1, \cdots, L$, are zero mean Gaussian distributed, and independent from each other. Variance of δ_l is computed as:

$$\sigma_{\delta,l}^2 = \frac{1}{2} E\left|\frac{\underline{x}_l^h \underline{n}_l}{\underline{x}_l^h \underline{x}_l}\right|^2 = \frac{1}{2}\left(\frac{1}{\underline{x}_l^h \underline{x}_l}\right)^2 E\left(\underline{x}_l^h \underline{n}_l \underline{n}_l^h \underline{x}_l\right) = \frac{1}{\underline{x}_l^h \underline{x}_l} \sigma_{n,l}^2 , \tag{B5}$$

where in derivation of the last expression, we have used $\frac{1}{2}E\left(\underline{n}_l\underline{n}_l^h\right)=\sigma_{n,l}^2\mathbf{I}_{M_\alpha}$ that $\mathbf{I}_{M_\alpha}$ is the $M_\alpha \times M_\alpha$ identity matrix. In the IEEE 802.11 all OFDM training symbols, used for subchannels gain estimation, are equal: $x_{li}=x_l$, with $|x_l|=1$. In this case, equations (B3), and (B5) reduce to:

$$\hat{\alpha}_l=\frac{1}{M_\alpha x_l}\sum_{i=1}^{M_\alpha} y_{li}\ ,$$

$$\sigma_{\delta,l}^2=\frac{1}{M_\alpha}\sigma_{n,l}^2\quad . \tag{B6}$$

IX. REFERENCES

[1] LAN/MAN Specific Requirements – Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High Speed Physical Layer in the 5 GHz band, IEEE P802.11 Wireless LANs doc.: IEEE 802.11-98/370ar3, Nov. 1998.

[2] European Digital Cellular Telecommunications System (Phase 2): Cannel Coding (GSM 05.03), ETSI, 1995

[3] E. Malkamäki, and H. Leib, "Coded Diversity on Block-Fading Channels," *IEEE Trans. Inform. Theory*, Vol. 45, No. 2, March 1992, pp. 771-781.

[4] E. Malkamäki, and H. Leib, "Rate 1/n Convolutional Codes with Interleaving Depth of n over a Block Fading Rician Channel," *in Proc. IEEE Veh. Technol. Conf.*, Phonix, AZ, May 1997, pp. 2002-2006.

[5] R. Knop, and P.A. Humblet, "Maximizing Diversity on Block Fading Channels," *in Proc. IEEE Int. Conf. Commun. 97*, Montreal, Quebec, Canada, June 1997, pp. 647-651.

[6] A. R. S. Bahai, and B. Saltzberg, "Multi-Carrier Digital Communications" 1999, Kluwer, NY.

[7] E. Zehavi, “8-PSK Trellis Codes for a Rayleigh Channel,” *IEEE Trans. On Commun.,* Vol. 40, No. 5, pp. 873-884, May. 1992.

[8] G. Caire, G. Taricco, E. Biglieri, “Bit-Interleaved Coded Modulation,” *IEEE Trans. Inform. Theory,* Vol. 44, No. 3, pp. 927-946, May. 1998.

[9] X. Li, J.A. Ritcey, “Bit-Interleaved Coded Modulation with Iterative Decoding using Soft Feedback,” *Electronics Letters*, Vol. 34, No. 10, pp. 942-943, May 1998.

[10] R. Hoshyar, S.H. Jamali, and A. R. Bahai, “Iterative Decoding of Bit Interleaved Trellis Coded Modulation”, Turbo Coding Group (TCG) progress report TCG3, May 29, 1999.

[11] A. J. Viterbi, “Principles of Digital Communication and Coding,” McGraw-Hill, Inc., 1979.

[12] R. Hoshyar, S. H. Jamali, and A. R. Bahai, “Turbo Coding Performance in OFDM Packet Transmission,” 51st IEEE Vehicular Technology Conference Proceedings, VTC 2000-Spring Tokyo., pp. 805 –810, 2000

[13] M.K. Simon, J.K. Omura, R.A. Schltz, and B.K. Levitt, *Spread Spectrum Communications*. Vol. 1, Rockvile, Maryland: Computer Science Press, 1985.

[14] J. Hagenauer, E. Offer, and L. Papke, “Iterative Decoding of Binary Block and Convolutional Codes*,” IEEE Trans. Inform. Theory,* Vol. 42, No. 2, pp. 429-445, Mar. 1996.