

Downlink Transmission of Broadband OFCDM Systems—Part IV: Soft Decision

Yiqing Zhou, *Member, IEEE*, and Jiangzhou Wang, *Senior Member, IEEE*

Abstract—In this paper, the performance of turbo-coded orthogonal frequency and code-division multiplexing (OFCDM) systems is investigated with soft multicode interference (MCI) cancellation and minimum mean-square error (MMSE) detection for downlink transmission in future high-speed wireless communications. To regenerate the soft interference signal, the conventional turbo decoding algorithm must be modified to provide log-likelihood ratio (LLR) values for all coded bits. Based on the LLR outputs of turbo decoder, two soft-decision functions are proposed, called *LLR-soft-decision* and *Gaussian-soft-decision* functions. The Gaussian assumptions used for deriving these two soft functions are verified by simulation results, and simple methods are proposed to estimate parameters used in the soft functions in practical systems. By means of computer simulations, the performance of soft MCI cancellation is studied extensively and compared with that of hard ones. It is shown that in a highly frequency-selective channel, soft MCI cancellation and MMSE detection can significantly improve the performance of turbo-coded OFCDM systems. Two iterations in turbo decoding are sufficient for both hard and soft-decision functions. The proposed soft-decision functions outperform the hard-decision function with various channel conditions and system parameters, such as the channel correlation, the quality of channel estimation, the number of iterations in turbo decoding and the frequency-domain spreading factor (N_F). Furthermore, the Gaussian-soft-decision function provides better performance than the LLR-soft-decision function. Finally, although frequency diversity gain is saturated for large channel correlation when N_F is large as in [6], the gain increases further with increasing N_F for small channel correlation even when N_F is large.

Index Terms—Fading channel, interference cancellation, minimum mean-square error (MMSE) detection, orthogonal frequency-division multiplexing (OFDM), packet-error rate (PER), soft-decision function, two-dimensional (2-D) spreading.

I. INTRODUCTION

IN FUTURE wireless communications, it is widely believed that the data rate on the downlink transmission should be increased significantly for a system to provide advanced multimedia services such as watching movie in real time. Therefore, a huge bandwidth (e.g., 100 MHz [1]) is needed in downlink. Since the multipath interference is severe in such a broadband channel, the orthogonal frequency-division multiplexing (OFDM) technique has been considered due to its robustness to multipath interference. Moreover, code spreading is introduced in the OFDM system to provide more flexibility

to the system deployment. Combining the OFDM with two-dimensional (2-D) code spreading, which includes time and frequency-domain spreading, the orthogonal frequency and code-division multiplexing (OFCDM) system has been proposed as a promising candidate for the downlink transmission in future broadband wireless communications [2], [3].

In the OFCDM systems, the orthogonal variable spreading factor (OVSF) codes [4] are employed as the 2-D spreading codes. Each data symbol is spread in time-domain with N_T chips and in frequency-domain with N_F chips. It should be noted that the N_F subcarriers carrying the chips in frequency-domain should be separated as far as possible, so that the N_F subcarriers experience different fading and the maximum frequency diversity gain can be obtained. It is assumed that in downlink transmission, the system employs packet switch instead of circuit switch. Therefore, at any time, only one user is concerned and there is no multiple-user interference. Given the total spreading factor $N = N_T \times N_F$, a user data rate can be increased by using multicode transmission. However, in a realistic broadband mobile channel, the code orthogonality will be distorted by the possible fast fading and more seriously, the frequency selectivity. Therefore, in the OFCDM system, multicode interference (MCI) is present in the high-speed transmission. In order to improve the system performance, MCI cancellation must be considered.

The basic idea of MCI cancellation is to use the decision signals of previous stage to regenerate MCI, then subtract MCI from the received signal. The decision function could be hard or soft. In [5], the hybrid MCI cancellation and MMSE detection was proposed, where the hard decisions of the demodulation output were used. In [6], hard decisions are used for coded OFCDM systems. Compared with the pure MMSE, the hybrid detection can enhance the system performance significantly. The hard-decision function is easy to realize and the MCI can be totally eliminated if the tentative decisions of previous stage are correct. However, when the tentative decisions are wrong, the MCI will be increased considerably from error propagation, actually doubled for binary phase-shift keying (BPSK) modulation. In order to improve the system performance with hard decisions, a soft-decision function [7] should be used to regenerate MCI. The soft MCI cancellation approach is more complex, but it can provide better performance over the hard one by alleviating the effect of error propagation.

In OFCDM systems, since the spread chips in frequency-domain experience different fading in a highly frequency-selective broadband channel, efficient signal combining techniques must be employed in the 2-D despreading. In this paper, a slow fading and highly frequency-selective channel is considered. Equal

Manuscript received April 1, 2005; revised October 1, 2005.

Y. Zhou is with Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong (e-mail: yqzhou@eee.hku.hk).

J. Wang is with Department of Electronics, University of Kent, Kent CT2 7NT, U.K. (e-mail: j.z.wang@kent.ac.uk).

Digital Object Identifier 10.1109/JSAC.2005.864010

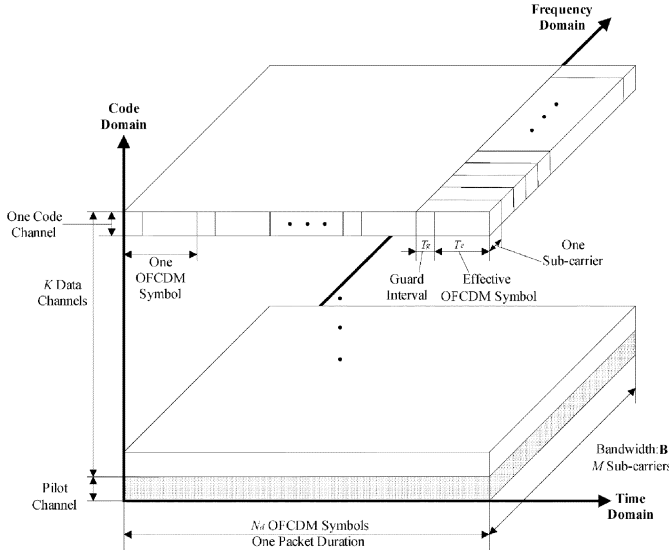


Fig. 1. Three-dimensional packet structure of the OFCDM system.

gain combining (EGC) is employed with time despreading, while minimum mean-square error (MMSE) combining is used in frequency despreading. The objective of this paper is to investigate the performance of soft MCI cancellation in turbo-coded OFCDM systems with the combining techniques.

In the concerned OFCDM system with quadrature phase-shift keying (QPSK) modulation and turbo coding, soft data decisions are carried out on the decoding output. Since all coded bits are required to regenerate the MCI, a modified turbo decoding algorithm is proposed to provide the log-likelihood ratio (LLR) of both systematic and parity bits. Based on the LLR output, the optimal soft-decision functions which minimize the decision errors are derived. Moreover, practical algorithms are also proposed to estimate the MMSE weighting factors and the parameters in the soft-decision functions.

The rest of this paper is organized as follows. Section II describes the OFCDM system in detail, including the packet structure, transmitter, channel model, and receiver. The operation of soft MCI cancellation and MMSE detection is shown in Section III, which describes the modified turbo decoding algorithm and derives the optimum soft-decision functions. Then, the performance of the turbo-coded OFCDM system with soft MCI cancellation and MMSE detection is evaluated in Section IV by computer simulations. Finally, conclusions are drawn in Section V.

II. SYSTEM DESCRIPTION

A. Packet Structure

Fig. 1 illustrates the packet structure of the OFCDM system in three dimensions. In code domain, there are totally K data channels and one pilot channel, sharing the time and frequency resources. In order to obtain good channel estimation, the code-multiplexed pilot channel should be orthogonal to all data channels in time-domain. Therefore, K will be at most $(N_T - 1)N_F$. With the spreading factor $N = N_T \times N_F$, the k th code channel is spread by a dedicated 2-D spreading code $\{C_{N_T}^{(k_T)}, C_{N_F}^{(k_F)}\}$, where the time-do-

main spreading code, $C_{N_T}^{(k_T)} = \{c_{N_T,0}^{(k_T)}, \dots, c_{N_T,N_T-1}^{(k_T)}\}$, is an OVFS code with a length of N_T , and the frequency-domain spreading code, $C_{N_F}^{(k_F)} = \{c_{N_F,0}^{(k_F)}, \dots, c_{N_F,N_F-1}^{(k_F)}\}$, is an OVFS code with a length of N_F . $C_{N_T}^{(k_T)}$ and $C_{N_F}^{(k_F)}$ are the (k_T) th and (k_F) th codes on the $\log_2 N_T$ and $\log_2 N_F$ layers in the OVFS code tree [4], respectively. Suppose that the zeroth time-domain spreading code with all “+1” is assigned to the pilot channel. The relationship among k , k_T and k_F is $k_T = k - \lfloor k/(N_T - 1) \rfloor (N_T - 1) + 1$ and $k_F = \lfloor k/(N_T - 1) \rfloor$, where $k_T = 1, \dots, N_T - 1$, $k_F = 0, 1, \dots, N_F - 1$, and $\lfloor x \rfloor$ is the integer portion of x .

In frequency-domain, the whole bandwidth B is divided into M subbands with M subcarriers. Given the frequency-domain spreading factor, N_F , each data code channel can transmit $N_B = M/N_F$ symbols on the M subcarriers in parallel. Finally, in time-domain, each packet is composed of N_d OFCDM symbols, each one consisting of a guard interval of T_g and an effective OFCDM symbol of T_e . The guard interval is inserted between the effective OFCDM symbols to prevent intersymbol interference in a multipath channel. In one packet duration, each data code channel can transmit $M_D = N_d/N_T$ unspread symbols per subcarrier. In summary, totally $KN_B M_D$ data symbols are contained in one packet.

B. Transmitter

The basic structure of the turbo-coded OFCDM system is illustrated in Fig. 2. It is assumed that the \bar{k} th data channel with the 2-D spreading code $\{C_{N_T}^{(\bar{k}_T)}, C_{N_F}^{(\bar{k}_F)}\}$ is the desired channel. The data bits are first turbo-encoded [8], and then translated into quadrature phase-shift keying (QPSK) symbols by the modulator. A symbol interleaver is employed to improve the coding gain by randomizing the burst channel errors. The interleaved data symbols are then serial-to-parallel (S/P) converted into N_B streams, modulating M subcarriers after spreading. Since the processing of all data symbols is similar, Fig. 2 illustrates the operation on one data symbol as an example. As shown in Fig. 2, each symbol is first spread into N_T chips in time-domain with the spreading code $C_{N_T}^{(\bar{k}_T)} = \{c_{N_T,0}^{(\bar{k}_T)}, \dots, c_{N_T,N_T-1}^{(\bar{k}_T)}\}$. Then, the time-domain spread signal is duplicated into N_F copies and multiplied by the frequency-domain spreading code $C_{N_F}^{(\bar{k}_F)} = \{c_{N_F,0}^{(\bar{k}_F)}, \dots, c_{N_F,N_F-1}^{(\bar{k}_F)}\}$. Therefore, totally $N = N_T \times N_F$ spread chips per data symbol are obtained by means of the 2-D spreading. The other $N_B - 1$ data symbols are processed in the same way. Hence, totally $M = N_B N_F$ spread chips are obtained in frequency-domain. Before being up-converted to subcarriers, these M chips are interleaved so that the N_F subcarriers corresponding to the N_F chips carrying the same data symbol are separated in the frequency band, experiencing different fading and providing maximum frequency diversity gain. Using a uniform block interleaver, the zeroth data symbol uses the 0th, (N_B) th, $(2N_B)$ th, \dots , $((N_F - 1)N_B)$ th subcarrier, the first data symbol uses the 1st, $(N_B + 1)$ th, \dots ,

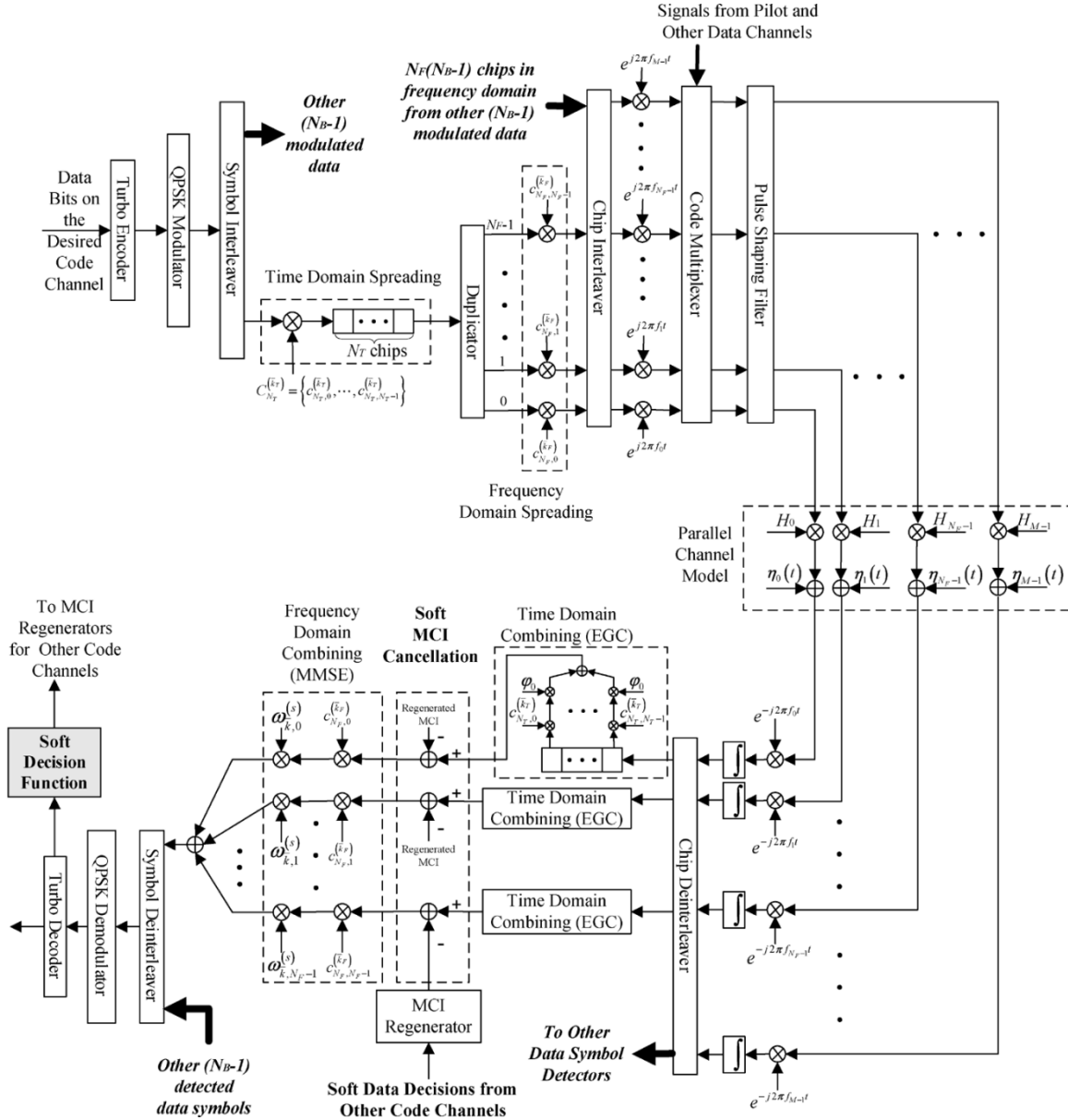


Fig. 2. Basic structure of the OFCDM system.

$((N_F - 1)N_B + 1)$ th subcarrier, and so on. After interleaving, the M chips will be up-converted to M subcarriers and transmitted in parallel. Note that this up-conversion at the transmitter and the down-conversion at the receiver could be realized by an M -point IFFT and FFT block, respectively, in practical systems. The signals from the K code channels and the pilot channel will be added together at the code multiplexer. The transmitted signal is given by (1) shown at the bottom of the page, where P is the signal power of one data code channel on one subcarrier, β is the power ratio of the pilot channel to one data channel, $d_{k,m,i}$ is the modulated data symbol with

unity power of the k th data channel on the m th subcarrier in the i th time chip (or OFCDM symbol) duration, d_P is the known QPSK symbol in the pilot channel with the all “+1” spreading code, f_m is the baseband equivalent frequency of the m th subcarrier, and $p(t)$ is the rectangular pulse shaping filter.

C. Channel Model

Under the assumption of perfect suppression of multipath by the guard interval, a parallel slow fading multichannel model is employed [9]. The signal transmitted on each subcarrier is multiplied by a complex fading factor H_m ($m = 0, \dots, M - 1$),

$$S(t) = \sum_{i=0}^{N_d-1} \sum_{m=0}^{M-1} \left(\sum_{k=0}^{K-1} \sqrt{P} d_{k,m,i} c_{N_T,i}^{(k_T)} c_{N_F,m}^{(k_F)} + \sqrt{\beta P} d_P \right) e^{j2\pi f_m(t-i(T_e+T_g))} p(t-i(T_e+T_g)) \quad (1)$$

then added by a complex additive white Gaussian noise (AWGN) $\eta_m(t)$ with a power spectral density of N_0 . The amplitude and phase of H_m are assumed to be Rayleigh distributed with $E\{|H_m|^2\} = 1$ and uniformly distributed in $[0, 2\pi]$, respectively. The correlation between any two channel fades H_{m_1} and H_{m_2} is introduced by the frequency separation $\Delta f_{m_1, m_2} = f_{m_1} - f_{m_2}$, given by [10], [11]

$$\rho_{m_1, m_2} = E\{H_{m_1}^* H_{m_2}\} = \frac{1}{\sqrt{1 + \left(\frac{\Delta f_{m_1, m_2}}{\Delta f_c}\right)^2}} \quad (2)$$

where $(\cdot)^*$ stands for the conjugate operation and Δf_c is the channel coherence bandwidth. The correlated Rayleigh-fading channels can be generated according to [12].

D. Receiver

After passing through the parallel multichannel, the received signal is first down-converted, and then passes through the integrator. The resultant signal on the m th subcarrier in the i th OFCDM symbol duration is given by

$$r_{m,i} = \left(\sum_{k=0}^{K-1} \sqrt{P} d_{k,m,i} c_{N_T,i}^{(k_T)} c_{N_F,m}^{(k_F)} + \sqrt{\beta P} d_P \right) H_m + \eta_{m,i} \quad (3)$$

where the variance of the noise term $\eta_{m,i}$ is N_0/T_e .

In a slow-fading channel, using the orthogonally code multiplexed pilot channel, the channel fading on each subcarrier can be estimated by first despreading the pilot channel, and then averaging the signals over the whole packet duration. Since the orthogonality in time-domain between the pilot channel and data channels preserves at the receiver side, the estimated channel fading is interference-free and only corrupted by the channel noise. The estimated channel fading is given by

$$\hat{H}_m = H_m + \varepsilon_m \quad (4)$$

where the estimation noise ε_m is a zero-mean complex Gaussian distributed variable with a variance of $N_0/(\beta P N_d T_e)$. With the estimated fading factors, the signals are then combined in time-domain, where EGC combining is employed. The weight of EGC on the m th subcarrier is given by

$$\varphi_m = \frac{\hat{H}_m^*}{|\hat{H}_m|}. \quad (5)$$

It should be noted that except the desired code $\{C_{N_T}^{(\bar{k}_T)}, C_{N_F}^{(\bar{k}_F)}\}$, the rest $(K-1)$ 2-D codes can be divided into two subsets: $\Omega_F = \left\{ \left\{ C_{N_T}^{(\bar{k}_T)}, C_{N_F}^{(k_F)} \right\} \middle| k_F \neq \bar{k}_F \right\}$ and $\Omega_T = \left\{ \left\{ C_{N_T}^{(k_T)}, C_{N_F}^{(\bar{k}_F)} \right\} \middle| k_T \neq \bar{k}_T \right\}$. After passing through a slow fading channel, the orthogonality between the desired code and the code channels in Ω_T can be preserved. Thus, the output of the time-domain combiner is only interfered by signals from the code channels in Ω_F . On the m th subcarrier, after time-domain despreading with EGC combining of the

zeroth to $(N_T - 1)^{\text{th}}$ OFCDM symbol, the resultant signal of the \bar{k} th code channel is given by

$$\begin{aligned} r_{\bar{k},m,0} &= \frac{1}{N_T} \sum_{k=0}^{K-1} \sum_{i=0}^{N_T-1} \sqrt{P} d_{k,m,i} c_{N_T,i}^{(k_T)} c_{N_F,m}^{(k_F)} H_m c_{N_T,i}^{(\bar{k}_T)} \varphi_m \\ &\quad + \frac{1}{N_T} \sum_{i=0}^{N_T-1} \sqrt{\beta P} d_P H_m c_{N_T,i}^{(\bar{k}_T)} \varphi_m + \eta_{\bar{k},m} \\ &= \sqrt{P} d_{\bar{k},m,0} c_{N_F,m}^{(\bar{k}_F)} (H_m \varphi_m) \\ &\quad + \sqrt{P} \sum_{k \in \Omega_F} d_{k,m,0} c_{N_F,m}^{(k_F)} (H_m \varphi_m) + \eta_{\bar{k},m} \end{aligned} \quad (6)$$

where $r_{\bar{k},m,0}$ corresponds to the zeroth data symbol of the M_D symbols in time-domain, and the noise term $\eta_{\bar{k},m}$ has a variance of $N_0/(N_T T_e)$. Note that the contribution from the pilot to the output of the time-domain combiner is zero. After the time-domain despreader, soft MCI cancellation is carried out. The re-generated soft MCI for the interested data signal on the desired code channel is given by

$$Q_{\bar{k},m,0}^{(s)} = \sqrt{P} \sum_{k \in \Omega_F} d_{k,m,0}^{(s)} c_{N_F,m}^{(k_F)} |\hat{H}_m|. \quad (7a)$$

It is subtracted from the output of the time-domain combiner, given by

$$\begin{aligned} r_{\bar{k},m,0}^{(s)} &= r_{\bar{k},m,0} - Q_{\bar{k},m,0}^{(s)} = \sqrt{P} d_{\bar{k},m,0} c_{N_F,m}^{(\bar{k}_F)} (H_m \varphi_m) \\ &\quad + \sqrt{P} \sum_{k \in \Omega_F} \left(d_{k,m,0} H_m \varphi_m - d_{k,m,0}^{(s)} |\hat{H}_m| \right) c_{N_F,m}^{(k_F)} + \eta_{\bar{k},m} \end{aligned} \quad (7)$$

where the superscript of s stands for the s th stage, and $d_{\bar{k},m,0}^{(s)}$ is the soft data decision. The resultant signals are then combined in frequency-domain with the MMSE technique. Since the weight of MMSE is related to the input interference power, it should be updated stage by stage as the MCI cancellation goes on. The MMSE weight at the s th stage is given by

$$\omega_{\bar{k},m}^{(s)} = \frac{E \left\{ \sqrt{P} d_{\bar{k},m,0} \left(r_{\bar{k},m,0}^{(s)} \right)^* \right\} c_{N_F,m}^{(\bar{k}_F)}}{E \left\{ \left| r_{\bar{k},m,0}^{(s)} \right|^2 \right\}} = \frac{P H_m^* \varphi_m^*}{E \left\{ \left| r_{\bar{k},m,0}^{(s)} \right|^2 \right\}}. \quad (8)$$

Consider the frequency-domain despreading with MMSE combining of the 0 th, $(N_B)th$, $(2N_B)th, \dots$, $((N_F - 1)N_B)th$ subcarriers. The output of the MMSE combiner for the desired code channel is given by (9) shown at the bottom of the next page, where $y_{\bar{k},0,0}^{(s)}$ corresponds to the zeroth data symbol of the N_B symbols in frequency-domain and the zeroth data symbol of the M_D symbols in time-domain. Equation (9) is composed of the useful signal component, noise term $\eta_{\bar{k},0,0}$, and the residual MCI $MCI_{\bar{k},0,0}^{(s)}$. It can be seen from (9) that the output of the MMSE detection is biased by a factor of $\sum_{m=0}^{N_F-1} (H_m N_B \varphi_m N_B) \omega_{\bar{k},m N_B}^{(s)}$. Therefore, $y_{\bar{k},0,0}^{(s)}$ should be normalized by the factor to get the unbiased signal. After normalization, the complex resultant signals are then demodulated and turbo decoded to recover the transmitted information bit sequence.

III. SOFT MCI CANCELLATION AND MMSE DETECTION

A. Estimation of MMSE Weights

In order to carry out MMSE detection and get the bias factor in practical systems, $\omega_{k,m}^{(s)}$ [see (8)] must be estimated. First of all, given the data power P , with the channel estimation \hat{H}_m , the numerator of $\omega_{k,m}^{(s)}$ can be estimated as $P|\hat{H}_m|$. Second, since the data symbols transmitted on different code channels are independent of each other, after soft MCI cancellation, the signals on different code channels should have the same power, i.e., $E\left\{|r_{k,m,0}^{(s)}|^2\right\}$ is same for all k . Moreover, assuming that in a slow-fading channel, the signals on each subcarrier experience the same fading in a packet duration, $E\left\{|r_{k,m,j}^{(s)}|^2\right\}$ is same for $j = 0, \dots, M_D - 1$. Finally, since in frequency-domain, the fading factors on nearby subcarriers are highly correlated, $E\left\{|r_{k,m,j}^{(s)}|^2\right\}$ is approximately the same for $m, m \pm 1, \dots, m \pm I_{\text{Favg}}$, where I_{Favg} is the window size. When the coherence bandwidth of the frequency-selective channel is large, I_{Favg} can take a large value due to the high correlation. On the contrary, when the coherence bandwidth is small, the correlation between subcarriers is reduced and a small I_{Favg} is expected. In summary, the estimation of $E\left\{|r_{k,m,0}^{(s)}|^2\right\}$ can be obtained by averaging

$$E\left\{|r_{k,m,0}^{(s)}|^2\right\} \approx \frac{1}{KM_D(2I_{\text{Favg}}+1)} \sum_{k=0}^{K-1} \sum_{j=0}^{M_D-1} \sum_{m'=m-I_{\text{Favg}}}^{m+I_{\text{Favg}}} |r_{k,m',j}^{(s)}|^2 \quad (10)$$

which is an unbiased estimator. Therefore, $\omega_{k,m}^{(s)}$ can be estimated as

$$\omega_{k,m}^{(s)} \approx \frac{KM_D(2I_{\text{Favg}}+1)P|\hat{H}_m|}{\sum_{k=0}^{K-1} \sum_{j=0}^{M_D-1} \sum_{m'=m-I_{\text{Favg}}}^{m+I_{\text{Favg}}} |r_{k,m',j}^{(s)}|^2}. \quad (11)$$

Thus, the biased factor is approximated by $\sum_{m=0}^{N_F-1} (H_{mN_B} \varphi_{k,mN_B}) \omega_{k,mN_B}^{(s)} \approx \sum_{m=0}^{N_F-1} |\hat{H}_{mN_B}| \omega_{k,mN_B}^{(s)}$ and the normalized signal is given by

$$\bar{y}_{k,0,0}^{(s)} = \frac{\sum_{m=0}^{N_F-1} r_{k,mN_B,0}^{(s)} \omega_{k,mN_B}^{(s)} c_{N_F,m}^{(\bar{k}_F)}}{\sum_{m=0}^{N_F-1} |\hat{H}_{mN_B}| \omega_{k,mN_B}^{(s)}} = \sqrt{P} d_{k,0,0} + \xi_{k,0,0}^{(s)} \quad (12)$$

where the term $\xi_{k,0,0}^{(s)}$ includes the noise caused by factor estimation error, residual MCI, and channel noise.

B. Modified Max_Log_Map Turbo Decoder

It is assumed that the normalized detection output $\bar{y}_{k,n,j}^{(s)}$ is corrupted by a zero-mean complex Gaussian noise $\xi_{k,n,j}^{(s)}$, where $n = 0, \dots, N_B - 1$ and $j = 0, \dots, M_D - 1$. The in-phase and quadrature components of $\bar{y}_{k,n,j}^{(s)}$ are independent of each other. Moreover, the variance of $\xi_{k,n,j}^{(s)}$ is supposed to be the same for all time and frequency positions, i.e., $\sigma_{\xi_{k,n,j}^{(s)}}^2(s) = \sigma_{k,\xi}^2(s)$. Therefore, after the QPSK demodulator, the input signals to the turbo decoder can be taken as BPSK modulated bit sequences passing through an AWGN channel with a noise variance of $\sigma_{k,\xi}^2(s)/2$.

Consider the demodulated signals on the \bar{k} th data channel. They can be grouped into sequences $\{y_{s,l}\}$ and $\{y_{p,l}\}$ corresponding to the systematic bit stream $\{b_{s,l}\}$ and parity bit stream $\{b_{p,l}\}$, respectively, where $l = 0, \dots, L_{\text{int}} - 1$ and L_{int} is the sequence length. The conventional turbo decoders proposed in [8] can be used. For implementation purpose, the less complex Max_Log_Map [13] algorithm is employed. It should be noted that the conventional decoding algorithm only provides LLR for the systematic bits. However, the LLR of all coded bits are needed to regenerate the soft MCI. Therefore, the algorithm should be modified to provide the LLR for parity bits as well.

Denote the coded output at step l as $\{b_{s,l}, b_{p,l}\}$. The original definition for the LLR of the systematic or parity bit is given by

$$\begin{aligned} \Lambda(b_{s/p,l}) &= \log \frac{\Pr(b_{s/p,l} = 1 | \text{observation})}{\Pr(b_{s/p,l} = 0 | \text{observation})} = \log \frac{\sum_u \sum_{u'} \lambda_{s/p,l}^1(u, u')}{\sum_u \sum_{u'} \lambda_{s/p,l}^0(u, u')} \\ &= \log \frac{\sum_u \sum_{u'} \gamma_{s/p,1}(y_{s,l}, y_{p,l}, u, u') \cdot \alpha_{l-1}(u') \cdot \beta_l(u)}{\sum_u \sum_{u'} \gamma_{s/p,0}(y_{s,l}, y_{p,l}, u, u') \cdot \alpha_{l-1}(u') \cdot \beta_l(u)} \quad (13) \end{aligned}$$

$$\begin{aligned} y_{k,0,0}^{(s)} &= \sum_{m=0}^{N_F-1} r_{k,mN_B,0}^{(s)} c_{N_F,m}^{(\bar{k}_F)} \omega_{k,mN_B}^{(s)} \\ &= \sqrt{P} d_{k,0,0} \sum_{m=0}^{N_F-1} (H_{mN_B} \varphi_{mN_B}) \omega_{k,mN_B}^{(s)} + \eta_{k,0,0} \\ &\quad + \underbrace{\sqrt{P} \sum_{k \in \Omega_F} \left(d_{k,0,0} \left(\sum_{m=0}^{N_F-1} c_{N_F,m}^{(k_F)} c_{N_F,m}^{(\bar{k}_F)} (H_{mN_B} \varphi_{mN_B}) \omega_{k,mN_B}^{(s)} \right) - d_{k,0,0} \left(\sum_{m=0}^{N_F-1} c_{N_F,m}^{(k_F)} c_{N_F,m}^{(\bar{k}_F)} |\hat{H}_{mN_B}| \omega_{k,mN_B}^{(s)} \right) \right)}_{\text{MCI}_{k,0,0}^{(s)}} \quad (9) \end{aligned}$$

where u and u' are the states at step l and $l-1$, respectively, and the sub “ s/p ” stands for systematic or parity. Conditioned on all the observations, $\{(y_{s,0}, y_{p,0}), \dots, (y_{s,L_{\text{int}}-1}, y_{p,L_{\text{int}}-1})\}$, $\Pr(b_{s/p,l} = i | \text{observation})$ ($i = 0, 1$) is the probability that $b_{s,l}$ or $b_{p,l}$ is i , and $\lambda_{s/p,l}^i(u, u')$ is the joint probability that $b_{s/p,l} = i$, the state at step l is u , and the state at step $l-1$ is u' . $\Pr(b_{s/p,l} = i | \text{observation})$ can be obtained from the joint probability $\lambda_{s/p,l}^i(u, u')$ by summing over all possible u and u' . $\lambda_{s/p,l}^i(u, u')$ is further divided into three items, $\alpha_{l-1}(u')$, $\beta_l(u)$, and $\gamma_{s/p,i}(y_{s,l}, y_{p,l}, u, u')$. Conditioned on the observations from step 0 to $l-1$, $\alpha_{l-1}(u')$, called as *the forward accumulated path metrics*, is the probability that the state at step $l-1$ is u' . $\beta_l(u)$, called as *the backward accumulated path metrics*, is a term related to the probability that the state at step l is u conditioned on the observations from step l to $L_{\text{int}}-1$. Finally, $\gamma_{s/p,i}(y_{s,l}, y_{p,l}, u, u')$, called as *the branch metrics*, is the joint probability that the state at step l is u , the coded output $b_{s/p,l} = i$, and the observed signal is $\{y_{s,l}, y_{p,l}\}$, conditioned on the state at step $l-1$ being u' .

Using the approximation of $\log \sum_j \exp x_j \approx \max_j (x_j)$, the LLR of $b_{s,l}$ or $b_{p,l}$ can be calculated as

$$\Lambda(b_{s/p,l}) = \max_{(u,u')} \left\{ \bar{\gamma}_{s/p,l}(y_{s,l}, y_{p,l}, u, u') + \bar{\alpha}_{l-1}(u') + \bar{\beta}_l(u) \right\} - \max_{(u,u')} \left\{ \bar{\gamma}_{s/p,0}(y_{s,l}, y_{p,l}, u, u') + \bar{\alpha}_{l-1}(u') + \bar{\beta}_l(u) \right\} \quad (14)$$

where $\bar{\alpha}_{l-1}(u') = \log \alpha_{l-1}(u')$, $\bar{\beta}_l(u) = \log \beta_l(u)$, and $\bar{\gamma}_{s/p,i}(y_{s,l}, y_{p,l}, u, u') = \log \gamma_{s/p,i}(y_{s,l}, y_{p,l}, u, u')$. For $b_{s,l}$, $\bar{\gamma}_{s,i}(y_{s,l}, y_{p,l}, u, u')$ can be calculated as in [13]. However, for $b_{p,l}$, the calculation of $\bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u')$ is given by

$$\begin{aligned} \bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u') &= \log \Pr\{y_{p,l} | b_{p,l} = i, u, u'\} + \log \Pr\{y_{s,l} | b_{p,l} = i, u, u'\} \\ &\quad + \log \Pr\{u | u'\} \\ &= \frac{\sqrt{2P}}{\sigma_{k,\xi}^2(s)} y_{p,l}(2i-1) + \frac{\sqrt{2P}}{\sigma_{k,\xi}^2(s)} y_{s,l}(2b_{s,l}-1) + \log \Pr\{u | u'\}. \end{aligned} \quad (15)$$

A priori information $\log \Pr\{u | u'\}$ depends on the extrinsic information of the systematic bit $L(b_{s,l})$ provided by the previous decoder, given by (16) shown at the bottom of the page. By assuming that $L(b_{s,l})$ is proportional to $\sqrt{2P}/\sigma_{k,\xi}^2(s)$, $\bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u')$ is also proportional to this ratio. Therefore, $\sqrt{2P}/\sigma_{k,\xi}^2(s)$ can be omitted from the metric calculations,

and $\bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u')$ is given by (17) shown at the bottom of the page. With $\bar{\alpha}_{l-1}(u')$ and $\bar{\beta}_l(u)$ calculated for the LLR of systematic bits, and the modified branch metrics $\bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u')$, the LLR of parity bits $b_{p,l}$ can be obtained. Since the LLR of parity bits is only calculated in the last iteration, the excess complexity caused by the modified Max_Log_Map algorithm is limited.

C. Optimal Soft-Decision Functions

At the output of the turbo decoder, the LLRs of systematic bits and parity bits are combined into one LLR stream, which is denoted as $\Lambda_{k,l}^{(s)}$, $l = 0, 1, \dots, 2M_D N_B - 1$ for the k th data channel at the s th stage, where $2M_D N_B$ is the number of coded bits transmitted on one data channel. The soft data decision is given by $b_{k,l}^{(s)} = f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)})$, where $f_{\text{dec}}^{(s)}(x)$ is the soft-decision function and can be found by minimizing the conditional mean-square decision error (MSDE) $E \left\{ \left| b_{k,l} - f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) \right|^2 \middle| \Lambda_{k,l}^{(s)} \right\}$ [14], [15], where $b_{k,l} = \pm \sqrt{P/2}$ is the transmitted coded bit before QPSK modulation. By differentiating the conditional MSDE with respect to $f_{\text{dec}}^{(s)}(x)$ and letting the differentiation be zero, one obtains

$$f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) = E \left\{ b_{k,l} \middle| \Lambda_{k,l}^{(s)} \right\} = \sum_{b_{k,l} = \pm \sqrt{P/2}} b_{k,l} \Pr \left\{ b_{k,l} \middle| \Lambda_{k,l}^{(s)} \right\}. \quad (18)$$

It can be seen from (18) that the soft-decision function $f_{\text{dec}}^{(s)}(x)$ is decided by the statistic property of the turbo decoding output $\Lambda_{k,l}^{(s)}$. Based on different interpretations of $\Lambda_{k,l}^{(s)}$, two different soft-decision functions can be obtained. First, the physical meaning of $\Lambda_{k,l}^{(s)}$ is exploited, i.e., $\Lambda_{k,l}^{(s)}$ is the LLR of coded bits. Using the definition of LLR, a soft-decision function called *LLR-soft-decision* function will be derived, where the probability density function (pdf) of $\Lambda_{k,l}^{(s)}$ is not needed at all. On the other hand, regardless of the original definition, the final turbo decoding output $\Lambda_{k,l}^{(s)}$ can be approximated as a BPSK modulated signal passing through an AWGN channel [8]. Based on this assumption, the pdf of $\Lambda_{k,l}^{(s)}$ can be obtained and another soft-decision function called *Gaussian-soft-decision* function will be derived.

1) *LLR-Soft-Decision Function*: First of all, the original definition of LLR is exploited to obtain $f_{\text{dec}}^{(s)}(x)$. In the

$$\log \Pr \{u | u'\} = \begin{cases} L(b_{s,l}) - \max(0, L(b_{s,l})), & \text{when } \Pr \{b_{s,l} = 1 | u', u, b_{p,l} = i\} = 1 \\ -\max(0, L(b_{s,l})), & \text{when } \Pr \{b_{s,l} = 0 | u', u, b_{p,l} = i\} = 1 \end{cases} \quad (16)$$

$$\bar{\gamma}_{p,i}(y_{s,l}, y_{p,l}, u, u') = \begin{cases} y_{p,l}(2i-1) + y_{s,l} + [L(b_{s,l}) - \max(0, L(b_{s,l}))], & \text{when } \Pr \{b_{s,l} = 1 | u, u', b_{p,l} = i\} = 1 \\ y_{p,l}(2i-1) - y_{s,l} + [-\max(0, L(b_{s,l}))], & \text{when } \Pr \{b_{s,l} = 0 | u, u', b_{p,l} = i\} = 1 \end{cases} \quad (17)$$

Max_Log_Map algorithm, the ratio $\sqrt{2P}/\sigma_{k,\xi}^2(s)$ is omitted. The decoder output $\Lambda_{k,l}^{(s)}$ is actually a scaled LLR, given by

$$\Lambda_{k,l}^{(s)} = \frac{\sigma_{k,\xi}^2(s)}{\sqrt{2P}} \cdot \log \frac{\Pr\left(b_{k,l} = \sqrt{\frac{P}{2}} \middle| \text{observation}\right)}{\Pr\left(b_{k,l} = -\sqrt{\frac{P}{2}} \middle| \text{observation}\right)}. \quad (19)$$

Since

$$\Pr\left(b_{k,l} = \sqrt{\frac{P}{2}} \middle| \text{observation}\right) + \Pr\left(b_{k,l} = -\sqrt{\frac{P}{2}} \middle| \text{observation}\right) = 1$$

$$\Pr\left(b_{k,l} = \sqrt{P/2} \middle| \text{observation}\right)$$

and

$$\Pr\left(b_{k,l} = -\sqrt{P/2} \middle| \text{observation}\right)$$

can be derived as

$$\begin{cases} \Pr\left(b_{k,l} = \sqrt{\frac{P}{2}} \middle| \text{observation}\right) = \frac{e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}}{1 + e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}} \\ \Pr\left(b_{k,l} = -\sqrt{\frac{P}{2}} \middle| \text{observation}\right) = \frac{1}{1 + e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}} \end{cases}. \quad (20)$$

Therefore, the optimal soft-decision function can be derived as

$$\begin{aligned} f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) &= \sum_{b_{k,l}} b_{k,l} \Pr\left\{b_{k,l} \middle| \Lambda_{k,l}^{(s)}\right\} \\ &= \sqrt{\frac{P}{2}} \cdot \frac{e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}}{1 + e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}} + \left(-\sqrt{\frac{P}{2}}\right) \cdot \frac{1}{1 + e^{\Lambda_{k,l}^{(s)} \sqrt{2P}/\sigma_{k,\xi}^2(s)}} \\ &= \sqrt{\frac{P}{2}} \tanh\left(\frac{\sqrt{2P}}{2\sigma_{k,\xi}^2(s)} \Lambda_{k,l}^{(s)}\right) \end{aligned} \quad (21)$$

where $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$. Note that since $\sigma_{k,\xi}^2(s)$ is updated stage by stage, the soft-decision functions are different at different stages.

Although $\sigma_{k,\xi}^2(s)$ is not required in the modified Max_Log_Map algorithm, it occurs in the optimum soft-decision function (21). Hence, the estimation of $\sigma_{k,\xi}^2(s)$ is needed. A data decision aided algorithm is proposed to estimate $\sigma_{k,\xi}^2(s)$. First, at the output of the MMSE detection, a hard decision is made on the observed complex signal sequence

$\{\bar{y}_{k,n,j}^{(s)}\}$ of the concerned code channel, denoted as $\{\bar{d}_{k,n,j}^{(s)}\}$. Then, the variance $\sigma_{k,\xi}^2(s)$ is given by

$$\sigma_{k,\xi}^2(s) = E\left\{\left|\xi_{k,n,j}^{(s)}\right|^2\right\} \approx \frac{1}{M_D N_B} \sum_{n=0}^{N_B-1} \sum_{j=0}^{M_D-1} \left|\bar{y}_{k,n,j}^{(s)} - \bar{d}_{k,n,j}^{(s)}\right|^2. \quad (22)$$

The accuracy of this estimator improves when the quality of channel estimation and MMSE detection output improves.

2) *Gaussian-Soft-Decision Function*: According to the research on turbo codes [8], the decoding output $\Lambda_{k,l}^{(s)}$ can be approximated as a BPSK modulated signal passing through an AWGN channel. The sign of the assumed BPSK signal is the same as that of $b_{k,l}$ and the amplitude is $E_{\Lambda,k}^{(s)}$. Therefore, $\Lambda_{k,l}^{(s)}$ can be approximated by

$$\Lambda_{k,l}^{(s)} = \frac{E_{\Lambda,k}^{(s)}}{\sqrt{\frac{P}{2}}} b_{k,l} + v_{k,l} \quad (23)$$

where $v_{k,l}$ is the assumed AWGN noise with a variance of $\sigma_{\Lambda,k}^2(s)$. Hence, the decision function $f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)})$ is given by

$$\begin{aligned} f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) &= \sum_{b_{k,l}} b_{k,l} \Pr\left\{b_{k,l} \middle| \Lambda_{k,l}^{(s)}\right\} \\ &= \sum_{b_{k,l}} b_{k,l} \frac{\Pr\left\{\Lambda_{k,l}^{(s)} \middle| b_{k,l}\right\} \Pr\left\{b_{k,l}\right\}}{\sum_{b_l} \Pr\left\{\Lambda_{k,l}^{(s)} \middle| b_{k,l}\right\} \Pr\left\{b_{k,l}\right\}} \\ &= \frac{\sum_{b_{k,l}} b_{k,l} e^{-\left(\Lambda_{k,l}^{(s)} - (E_{\Lambda,k}^{(s)}/\sqrt{P/2})b_{k,l}\right)^2 / (2\sigma_{\Lambda,k}^2(s))}}{\sum_{b_{k,l}} e^{-\left(\Lambda_{k,l}^{(s)} - (E_{\Lambda,k}^{(s)}/\sqrt{P/2})b_{k,l}\right)^2 / (2\sigma_{\Lambda,k}^2(s))}} \\ &= \sqrt{\frac{P}{2}} \tanh\left(\frac{E_{\Lambda,k}^{(s)}}{\sigma_{\Lambda,k}^2(s)} \Lambda_{k,l}^{(s)}\right). \end{aligned} \quad (24)$$

In practical systems, the parameters $E_{\Lambda,k}^{(s)}$ and $\sigma_{\Lambda,k}^2(s)$ can be estimated by (25) shown at the bottom of the page.

IV. SIMULATION RESULTS

A. Configuration

Computer simulation results are presented in this section to show the performance of soft MCI cancellation and MMSE detection in the turbo-coded OFCDM system. The major system parameters are defined as follows unless noted otherwise. The system employs a 2-D spreading factor of $N = N_T \times N_F = 4 \times 16$, and 48 code channels are used for

$$\begin{cases} E_{\Lambda,k}^{(s)} \approx E\left\{\left|\Lambda_{k,l}^{(s)}\right|\right\} \approx \frac{1}{2M_D N_B} \sum_{l=0}^{2M_D N_B-1} \left|\Lambda_{k,l}^{(s)}\right| \\ \sigma_{\Lambda,k}^2(s) \approx E\left\{\left|\Lambda_{k,l}^{(s)}\right|^2\right\} - \left[E\left\{\left|\Lambda_{k,l}^{(s)}\right|\right\}\right]^2 \approx \frac{1}{2M_D N_B} \sum_{l=0}^{2M_D N_B-1} \left|\Lambda_{k,l}^{(s)}\right|^2 - \left[E_{\Lambda,k}^{(s)}\right]^2 \end{cases} \quad (25)$$

data transmission. The power ratio between the pilot channel and all data channels β/K is set to 0.2. A bandwidth of 100 MHz is employed and totally $M = 1024$ subcarriers are used with a subcarrier spacing of 97.56 KHz. In one packet duration, there are $N_d = 48$ OFCDM symbols per subcarrier. Individually encoded by a turbo code [16] of 1/2 rate, the information bits transmitted on each data channel have a length of $(2M_D N_B - 12)/2$ when QPSK is considered, where 12 is the length of tail bits. Note that the length of information bits in turbo codes reduces when the total spreading factor N increases. Moreover, the prime interleaver [16] is employed as the internal interleaver of turbo codes. As to the parallel channel model, given a bandwidth of $B = 100$ MHz, the correlation between adjacent subcarriers carrying the same data symbol is set to $\rho_{0,64} = 0.075$ for $N_F = 16$. According to (2), the corresponding Δf_c can be obtained as 0.47 MHz. Other values of channel correlation will also be considered, corresponding to different Δf_c . The lowest $\rho_{0,64}$ considered in this paper is 0.05. In MMSE weight estimation, the optimal averaging window size $2I_{\text{Favg}} + 1$ in frequency-domain should be adjusted according to channel correlations. It is obtained from simulations that for $\rho_{0,64} = 0.05$ to 0.1, the optimal window size is three. Although the window size could be larger when higher channel correlation is considered, the performance difference between the optimal size and the size of three is not significant. Therefore, for simplicity reason, $2I_{\text{Favg}} + 1$ is set to three for all channel correlations. At the receiver side, the modified Max_Log_Map algorithm is employed in the turbo decoder. Finally, the effective information bit signal-to-noise ratio (SNR) is defined as $\text{SNR}_b = PN(1 + \beta/K)/2R_{\text{eff}}\sigma_n^2$, where 2 is the number of bits in one QPSK symbol and R_{eff} is the effective code rate which takes the tail bits in turbo codes into consideration.

B. Accuracy of Gaussian Assumptions for $\bar{y}_{k,n,j}^{(s)}$ and $\Lambda_{k,l}^{(s)}$

First of all, the Gaussian assumptions of the variables $\bar{y}_{k,n,j}^{(s)}$ and $\Lambda_{k,l}^{(s)}$ are verified by simulations. Note that $\bar{y}_{k,n,j}^{(s)}$ is the normalized MMSE detection output at the s th stage for the k th code channel, where n and j stand for the positions of the recovered data symbol in time and frequency domains, respectively. Given $P = 2$ and $\text{SNR}_b = 5.5$ dB, the number of iterations in turbo decoding is set to two. Using the system parameters defined before, totally $M_D N_B = 768$ QPSK symbols, or $2M_D N_B = 1536$ coded bits, equivalently, are transmitted on each data channel in one packet duration. The zeroth data code channel is of interest. Using the same set of channel fading factors, additive channel noise and data symbols are generated randomly. A seven-stage soft MCI cancellation is employed. When the LLR-soft-decision function is employed, the normalized MMSE detection outputs ($\bar{y}_{0,n,j}^{(s)}$) are recorded at each stage, while when the Gaussian-soft-decision function is used, the signals at the output of turbo decoder ($\Lambda_{0,l}^{(s)}$) are recorded. Totally, 100 packets are simulated. The parameters, i.e., $\sigma_{k,\xi}^2(s)$, $E_{\Lambda,k}^{(s)}$, and $\sigma_{\Lambda,k}^2(s)$, are estimated according to (22) and (25), respectively. The decoding output is then normalized by $E_{\Lambda,k}^{(s)}$. Histograms are shown in Fig. 3(a) and

(b) for the signals demodulated from all $\bar{y}_{0,n,j}^{(s)}$ in the zeroth and seventh stage, respectively, and in Fig. 3(c) for normalized decoded signals received in 100 packets. As a comparison, histograms are also plotted in Fig. 3 for Gaussian variables with the estimated variances of $\sigma_{k,\xi}^2(s)$ and $\sigma_{\Lambda,k}^2(s)$. In Fig. 3(a) and (b), for the histograms of the observed demodulated signals, it can be seen that at both the zeroth and seventh stages, the curves are concentrated on the possible transmitted signals, +1 and -1. At the zeroth stage, the MCI and noise in the demodulated signals are serious, and the curve is relatively flat. As soft MCI cancellation goes on, at the seventh stage, part of the MCI is cancelled out and the curve changes sharply. In both cases, the histograms of demodulated signals are similar to those obtained from Gaussian variables with the variance of $\sigma_{k,\xi}^2(s)$. Meanwhile, the histograms of the decoded signals in Fig. 3(c) show the same properties as those of the demodulated signals. However, compared with the demodulated signals, the decoded signals are much closer to the Gaussian variables. Hence, from Fig. 3, it can be concluded that the assumptions of Gaussian distribution for $\bar{y}_{k,n,j}^{(s)}$ and $\Lambda_{k,l}^{(s)}$ are reasonable, and the turbo decoder output $\Lambda_{k,l}^{(s)}$ can be more accurately modeled as Gaussian variables than $\bar{y}_{k,n,j}^{(s)}$.

C. System Performance

The packet-error rate (PER) and bit-error rate (BER) (after decoding) of the turbo-coded OFCDM system are illustrated in Fig. 4 when Gaussian-soft-decision MCI cancellation and MMSE detection is used. Since BER curves behave similarly to PER curves, only the description of PER is presented. It can be seen that at the zeroth stage without MCI cancellation, the system performance is quite poor. The PER performance is almost flat as SNR increases. Due to the low correlation $\rho_{0,64} = 0.075$, the MCI is serious and even the powerful turbo decoding cannot work in such a deteriorative environment. Although SNR increases, the system performance is limited by the MCI, and PER cannot be reduced. However, when soft MCI cancellation is introduced, the system performance can be enhanced considerably. The PER decreases as the number of stages increases due to the reduction of MCI in each stage, especially when SNR is large. This is because when MCI decreases, the system can benefit from the frequency diversity gain. Furthermore, turbo decoding begins to work and provides coding gain to the system. The most significant reduction in PER is obtained from the zeroth stage to the first stage, then the reduction becomes smaller as the number of stages increases. However, since the MCI in a channel with low correlation is severe, there is still considerable improvement in PER even at the seventh stage. For example, at $\text{SNR}_b = 5.5$ dB, the PER at the seventh stage is around 20% lower than that at the sixth stage. Therefore, in a highly frequency-selective channel, the performance of the turbo-coded OFCDM system can be improved significantly by the soft MCI cancellation. The larger the number of stages, the lower the PER. With $\rho_{0,64} = 0.075$, a seven-stage soft MCI cancellation can provide good performance.

Given $\text{SNR}_b = 5.5$ dB, using Gaussian-soft-decision MCI cancellation, the PER is plotted in Fig. 5 as a function of the

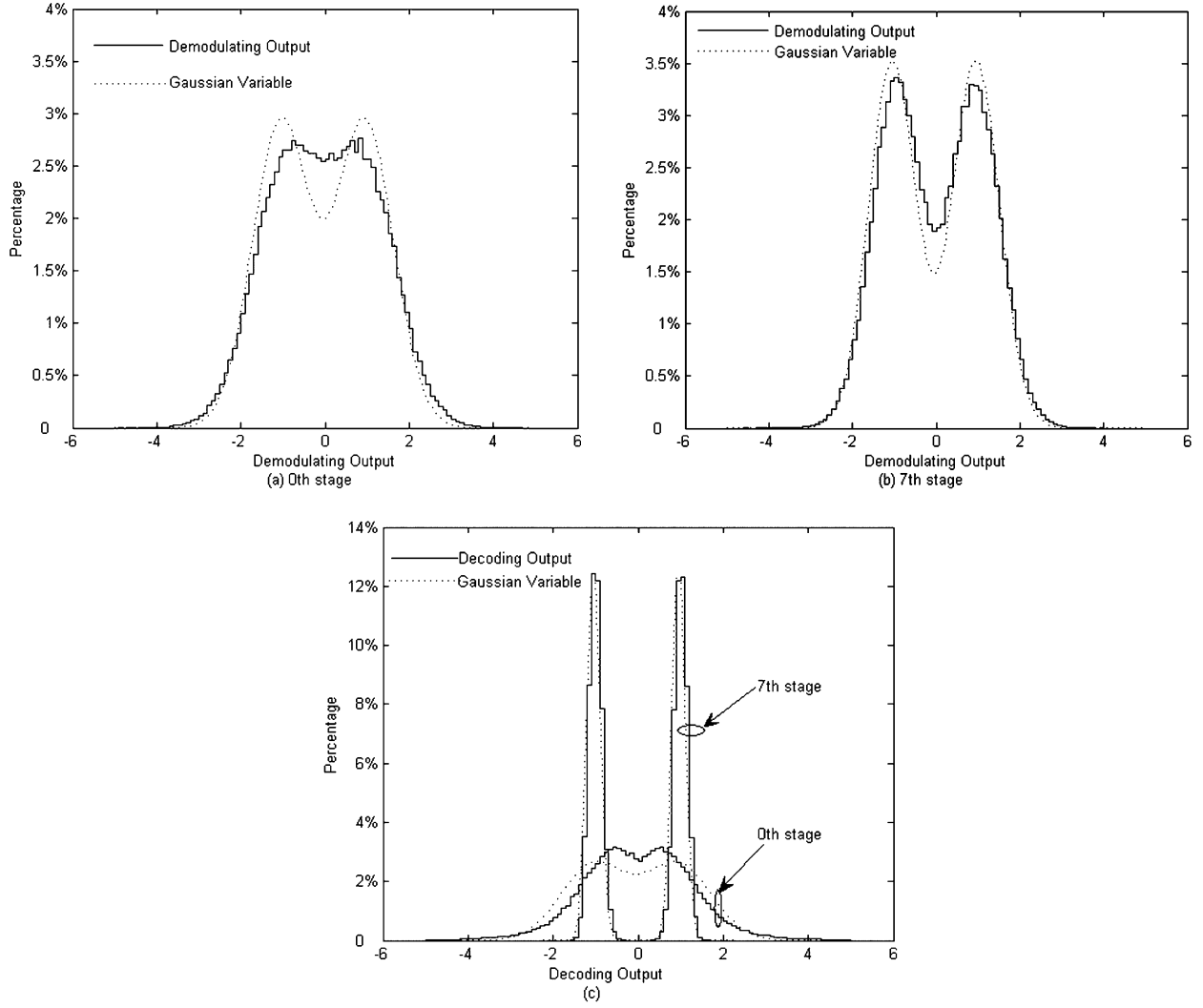


Fig. 3. Histograms of observed signals.

number of stages for various channel correlations. Note that Δf_c is varied in order to obtain different channel correlations. It is shown that with all kinds of correlations, the PER reduces as the number of stages increases. With a seven-stage soft MCI cancellation, the system in the lowest correlation obtains the most significant improvement in PER. On the other hand, when the channel correlation increases, there is less improvement in PER when the number of stages increases. It can be seen that for $\rho_{0,64} = 0.075$, a seven-stage soft MCI cancellation is needed. However, when $\rho_{0,64} = 0.2$, the PER reduction after the fourth stage is not significant. Hence, a four-stage soft MCI cancellation is sufficient for the system. Similarly, when $\rho_{0,64} = 0.4$, the performance improvement after the second stage is limited. Thus, a two-stage soft MCI cancellation is needed for $\rho_{0,64} = 0.4$. In conclusion, the number of stages required in the system depends on the channel correlations. If the system works in a channel with low correlations, a soft MCI cancellation with a large number of stages is preferred. However, if the concerned channel has high correlation, a small number of stages such as two or four is sufficient. A seven-stage MCI cancellation is employed in the following contents.

The system performance is further investigated in Fig. 6 as a function of channel correlation. Given $N_F = 16$, the channel correlation is represented by $\rho_{0,64}$. The PERs of the seventh stage in MCI cancellation are shown when hard-decision, LLR-soft-decision, and Gaussian-soft-decision functions are used. It can be seen that with MCI cancellation, the PERs become larger as the channel correlation $\rho_{0,64}$ increases. For various values of $\rho_{0,64}$, the Gaussian-soft-decision function always provides the best performance among the three decision functions, and the LLR-soft-decision function outperforms the hard-decision function. The superiority of Gaussian-over LLR-soft-decision function is because that the approximation used to obtain Gaussian-soft-decision function is more accurate than that used to obtain LLR-soft-decision function, as shown in Fig. 3. As a comparison, Fig. 6 also shows the system PER (after turbo decoding) at the zeroth stage without MCI cancellation. It can be seen that the PER at the zeroth stage slightly reduces as $\rho_{0,64}$ increases, which is contrary to the trend of the PERs with MCI cancellation. To explain the PER trends, the BERs before turbo decoding are shown in Fig. 6 for the zeroth and seventh stage when Gaussian-soft-decision is used. It is seen that the BERs

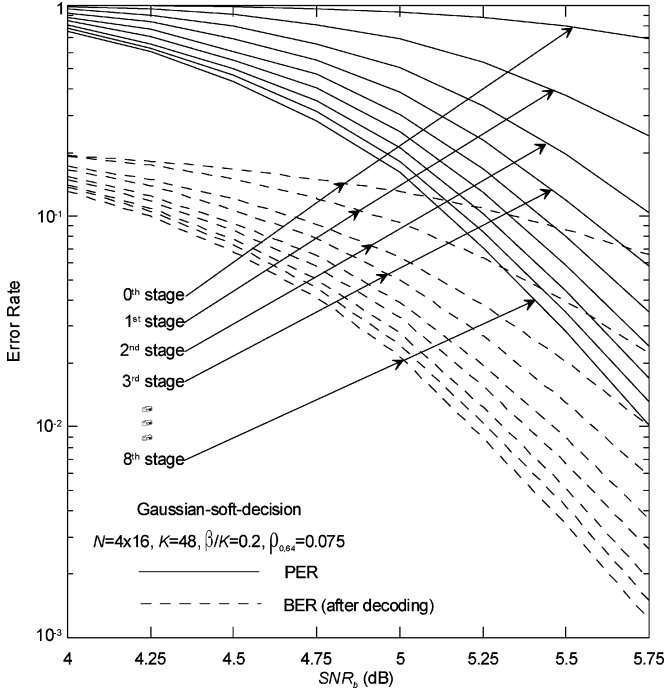


Fig. 4. Performance of the turbo-coded OFCDM system with soft MCI cancellation and MMSE detection.

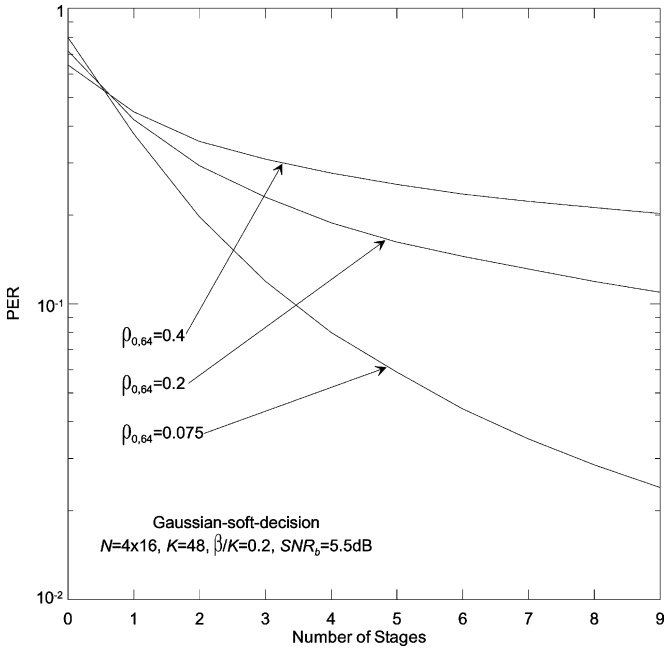


Fig. 5. PER performance as a function of the number of stages for various channel correlations.

increase with $\rho_{0,64}$. This is because although the MCI reduces as $\rho_{0,64}$ increases, the frequency diversity gain also decreases. As a result, the BER performance degrades. However, at the zeroth stage, the BER before turbo decoding is almost flat, so the change in input SNR to the decoder has little influence on the decoding output. On the other hand, the Max_Log_Map decoding algorithm is derived based on the assumption that the input signals have the same SNR. However, since the modulated symbols are spread with different set of subcarriers, the

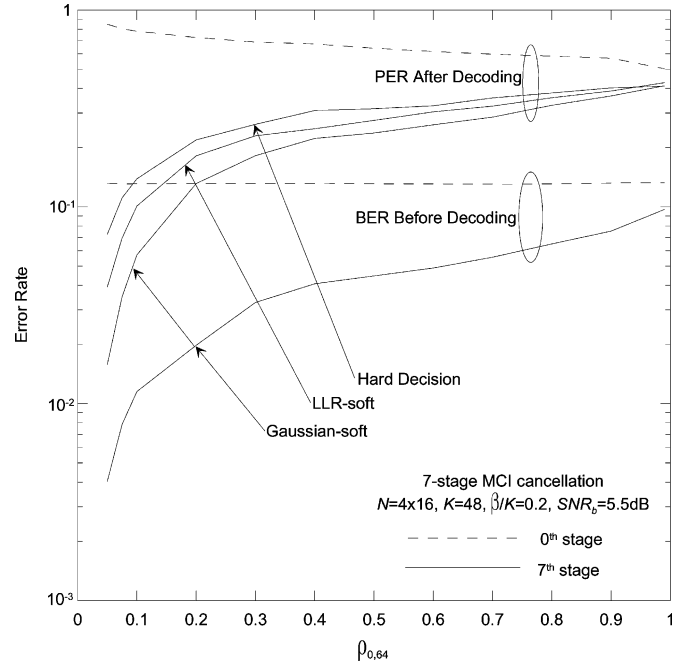


Fig. 6. System performance in channels with various correlations.

received SNR of each symbol is different. This SNR difference decreases when the channel correlation increases. Therefore, at the receiver side, the assumption that the inputs to the turbo decoder have the same SNR becomes more accurate as the channel correlation increases. As a result, the Max_Log_Map algorithm works better and the PER performance improves. On the other hand, at the seventh stage, the BER before turbo decoding increases rapidly with $\rho_{0,64}$. In this case, the input SNR dominates the performance of decoding algorithm and the PER degrades as $\rho_{0,64}$ increases.

Given $\rho_{0,64} = 0.075$ and $\text{SNR}_b = 5.5$ dB, Fig. 7 shows the PER performance comparison of the system with MCI cancellation using different decision functions. The soft-decision function with a fixed parameter, i.e., $f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) = \sqrt{P/2} \tanh(\alpha \cdot \Lambda_{l,k}^{(s)})$, is also investigated. The parameter α is changed from 0.01 to 5 in order to find its optimum value. It can be seen that when α is small, the PER is high and the performance is much worse than that of the hard decision. As α increases, the PER reduces and achieves a minimum value for α near 0.2. Choosing $\alpha \approx 0.2$, the soft-decision function $f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) = \sqrt{P/2} \tanh(\alpha \cdot \Lambda_{l,k}^{(s)})$ provides the best performance among all decision functions including the hard, LLR-soft and Gaussian-soft-decision functions. Then, as α is increased further, PER performance degrades, and is poorer than that of LLR-soft and Gaussian-soft-decision functions, and approaches to that of hard decision. Therefore, it can be seen that with properly chosen α , the soft-decision function $f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) = \sqrt{P/2} \tanh(\alpha \cdot \Lambda_{l,k}^{(s)})$ can achieve the best performance in all the investigated decision functions. This is because although LLR- and Gaussian-soft-decision functions are derived to minimize the MSDE, the derivations are based on some approximations of the real situations. Furthermore, the parameters used in LLR- and Gaussian-soft-decision functions need to be estimated, where

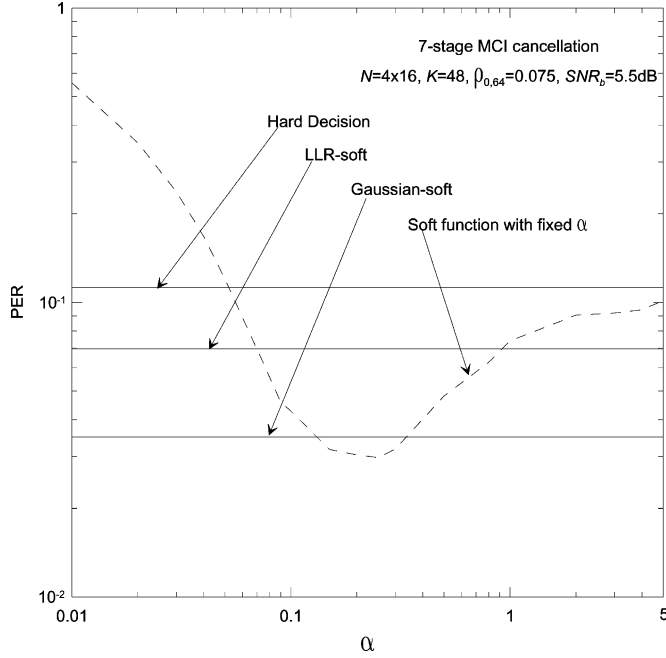


Fig. 7. Performance comparisons for different decision functions.

the estimation noise is present. Thus, it is possible that the function $f_{\text{dec}}^{(s)}(\Lambda_{k,l}^{(s)}) = \sqrt{P/2} \tanh(\alpha \cdot \Lambda_{k,l}^{(s)})$ outperforms the LLR- and Gaussian-soft-decision functions with appropriately chosen α . However, since the optimum α changes with channel conditions such as SNR and correlation, and the searching for the optimum α is time consuming, the soft-decision function with α is not applicable in practical systems. On the other hand, although the Gaussian-soft-decision function cannot achieve the minimum PER, its performance is stable and close to the optimal one. Furthermore, the estimation of parameters in the function [see (25)] is simple. Therefore, the Gaussian-soft-decision function is preferred in practical systems.

The effect of the pilot power ratio β/K on the system performance is illustrated in Fig. 8. For all the MCI cancellations with three different decision functions, when β/K is small, the PER is high due to the poor channel estimation. When β/K increases, more power is allocated to the pilot channel and better channel estimation is obtained. Therefore, PER performance improves as the quality of channel estimation improves, and PER reaches a minimum value for a particular value of β/K . Further increasing β/K beyond that value increases PER. This is because the total transmission power is limited. When too much power is assigned to the pilot channel, there is small power to data channels. So data channels become vulnerable to channel noise and the system performance degrades. In general, the power ratio β/K should be set in the range of [0.15, 0.25] to provide near-optimum system performance, irrespective of the decision functions. Moreover, for various values of β/K , the MCI cancellation using soft-decision functions always outperforms the one using hard-decision function, and the Gaussian-soft-decision function provides better performance than the LLR-soft-decision.

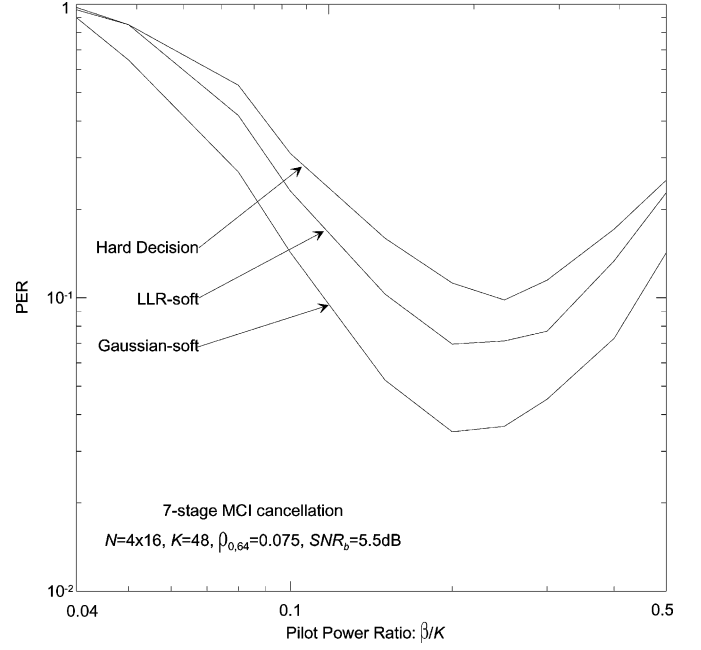
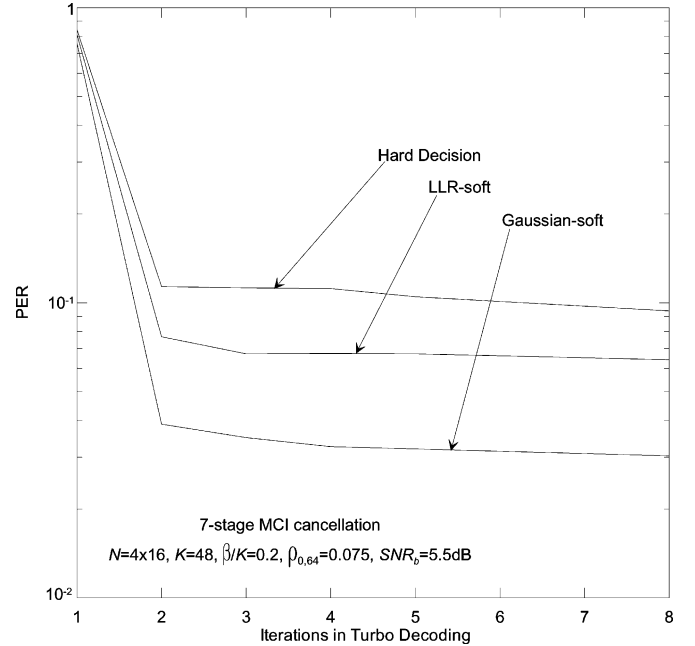
Fig. 8. System performance as a function of β/K .

Fig. 9. System performance as a function of the iterations in turbo decoding.

In the turbo-coded OFCDM system, the number of iterations used in turbo decoding is an important parameter. Fig. 9 shows the system PER as a function of the number of iterations. It can be seen that using a seven-stage MCI cancellation, a sharp reduction of PER is observed from the 1st iteration to the second iteration. Then, the PER curve tends to be flat after the second iteration. This is because for the hybrid detection with multiple stages, much MCI has been cancelled out, and only few iterations are needed. In summary, two iterations are sufficient to provide good performance for the turbo-coded OFCDM system with either hard or soft MCI cancellation.

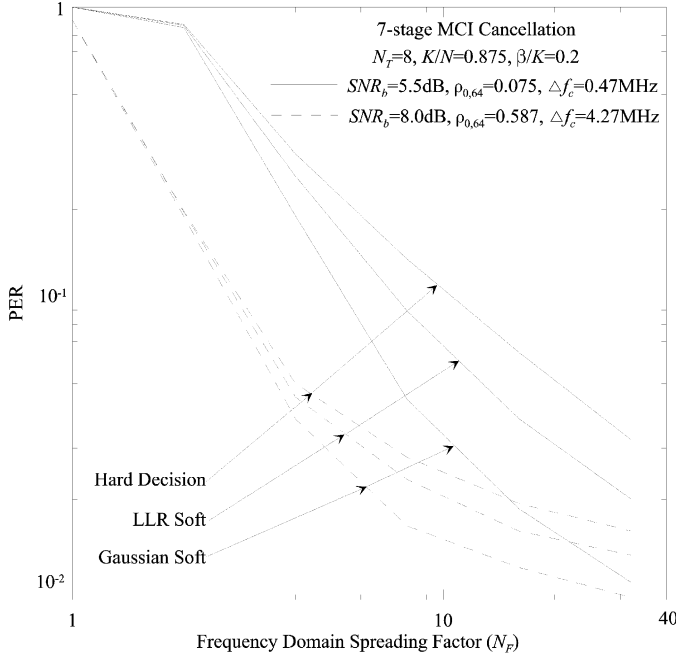


Fig. 10. PER performance as a function of frequency-domain spreading factor.

Given a time-domain spreading factor N_T of 8 and a system load of 0.875, the effect of the frequency-domain spreading factor, N_F , on the system performance is investigated in Fig. 10 for $SNR_b = 5.5$ dB and $\rho_{0,64} = 0.075$ ($\Delta f_c = 0.47$ MHz). For comparison, another set of curves (dashed) are also shown in Fig. 10 using similar parameters as that in [6]. Since a discrete multipath channel model is employed in [6], its frequency correlation is obtained by simply taking the Fourier transform of the delay power spectrum. Thus, the frequency correlation of the adjacent concerned subcarriers can be calculated as 0.564 for $N_F = 16$. The channel condition in [6] is approximated by the parallel multichannel model in the way that both channel models provide the same frequency correlation of the adjacent concerned subcarriers for $N_F = 16$. According to (2), the corresponding coherence bandwidth in the parallel multichannel model is $\Delta f_c = 4.27$ MHz. Then, given $SNR_b = 8.0$ dB, the system performance is evaluated using the parallel multichannel model with $\rho_{0,64} = 0.564$ ($\Delta f_c = 4.27$ MHz). For both small and large correlation, it can be seen that for $N_F = 1$, the system performance is not acceptable for either hard or soft-decisions. In spite of no MCI, the input to the turbo decoder is too noisy for the decoder to work due to no frequency diversity gain. When N_F is two, the PER reduces a little due to the frequency diversity gain. As N_F is increased further, the system performance improves significantly. This is because although MCI in frequency-domain increases, the frequency diversity gain also gets larger with increasing N_F . With MCI cancellation, the gain from diversity overcomes the loss from increased MCI. Therefore, the SNR of the input signal to turbo decoding improves, the decoder works effectively, and the system performance is enhanced. The frequency diversity gain is saturated for large correlation when N_F is large (i.e., $N_F \geq 16$). This is in consistency with [6]. However, the system performance can still be

improved significantly with increasing N_F for small channel correlation even when N_F is large. This is because for large correlation of $\rho_{0,64} = 0.564$, when N_F increases from 16 to 32, ρ_{0,N_B} increases from 0.564 to 0.807. Hence, the frequency diversity gain is saturated with the high ρ_{0,N_B} . However, for low correlation of $\rho_{0,64} = 0.075$, ρ_{0,N_B} is 0.15 for $N_F = 32$. With this low correlation and MCI cancellation, there is still considerable frequency diversity gain and the system performance can be improved significantly.

V. CONCLUSION

The performance of the turbo-coded OFCDM system has been investigated with soft MCI cancellation in this paper. Based on the likelihood output of turbo decoding, two soft-decision functions, i.e., LLR-soft-decision and Gaussian-soft-decision, are proposed. By means of computer simulation, the performance of MCI cancellation with soft- and hard-decision functions is studied extensively. The following conclusions are drawn.

- 1) Using Gaussian-soft-decision function as an example, soft MCI cancellation can improve the system performance significantly in a highly frequency-selective channel. Gaussian function outperforms LLR function. The number of stages needed in the MCI cancellation depends on the channel correlations. The lower the correlation is, the more the number of stages is needed.
- 2) In order to obtain near-optimum system performance, the power ratio of the pilot channel to all data channels should be taken in the range of [0.15, 0.25], irrespective of the decision functions used.
- 3) To carry out soft interference regeneration for the MCI cancellation, the conventional turbo decoding algorithm which only decodes systematic bits should be modified to decode parity bits as well. Furthermore, for both hard- and soft-decision functions, two iterations are sufficient in the turbo decoding to provide stable performance.
- 4) Although frequency diversity gain is saturated for large channel correlation when N_F is large as in [6], the gain increases further with increasing N_F for small channel correlation even N_F is large.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. Sawahashi and Dr. K. Higuchi from NTT DoCoMo for their helpful discussions and valuable comments.

REFERENCES

- [1] H. Atarashi and M. Sawahashi, "Investigation of inter-carrier interference due to Doppler spread in OFCDM broadband packet wireless access," *IEICE Trans. Commun. (Special Issue on Software Defined Radio Technologies and its Applications)*, vol. 86-B, no. 2002SRP-28, pp. 291–299, Jan. 2003.
- [2] H. Atarashi, S. Abeta, and M. Sawahashi, "Variable spreading factor-orthogonal frequency and code division multiplexing (VSF-OFCDM) for broadband packet wireless access," *IEICE Trans. Commun.*, vol. E86-B, pp. 291–299, Jan. 2003.

- [3] N. Maeda, Y. Kishiyama, and M. Sawahashi, "Variable spreading factor-OFCDM with two dimensional spreading that prioritizes time domain spreading for forward link broadband wireless access," in *Proc. IEEE Veh. Technol. Conf.-Spring*, Apr. 2003, pp. 127–132.
- [4] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Commun. Mag.*, vol. 36, no. 9, pp. 56–69, Sep. 1998.
- [5] Y. Zhou, J. Wang, and M. Sawahashi, "Downlink transmission of broadband OFCDM systems—Part I: Hybrid detection," *IEEE Trans. Commun.*, vol. COM-53, pp. 718–729, Apr. 2005.
- [6] —, "Downlink transmission of broadband OFCDM systems—Part III: Turbo-coded," *IEEE J. Sel. Areas Commun. (Special Issue on The Next Generation CDMA Technologies)*, vol. 24, no. 1, pp. 132–140, Jan. 2006.
- [7] R. Irmer, A. Nahler, and G. Fettweis, "On the impact of soft decision functions on the performance of multistage parallel interference cancellers for CDMA systems," in *IEEE Veh. Technol. Conf.-Spring*, vol. 2, May 2001, pp. 1513–1517.
- [8] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes (1)," in *Proc. IEEE ICC*, May 1993, pp. 1064–1070.
- [9] K. Witrisal, "OFDM air-interface design for multimedia communications," Ph.D. dissertation, Delft Univ. Technol., Delft, The Netherlands, 2002.
- [10] F. Lin and L. B. Milstein, "Successive interference cancellation in multicarrier DS/CDMA," *IEEE Trans. Commun.*, vol. 48, pp. 1530–1540, Sep. 2000.
- [11] W. C. Y. Lee, *Mobile Communications Engineering: Theory and Applications*, 2nd ed. New York: McGraw-Hill, 1997.
- [12] R. B. Ertel and J. H. Reed, "Generation of two equal power correlated Rayleigh fading envelopes," *IEEE Commun. Lett.*, vol. 2, pp. 276–278, Oct. 1998.
- [13] A. Worm, P. Hoeher, and N. Wehn, "Turbo-decoding without SNR estimation," *IEEE Commun. Letter*, vol. 4, pp. 193–195, Jun. 2000.
- [14] J. F. Rossler and J. B. Huber, "Iterative soft decision interference cancellation receivers for DS-CDMA downlink employing 4QAM and 16QAM," in *Proc. Conf. Record 36th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2002, pp. 1488–1494.
- [15] H. K. Jeong, S. J. Jung, and H. L. Yong, "On the use of sigmoid functions for multistage detection in asynchronous CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 522–526, Mar. 1999.
- [16] "Multiplexing and channel coding (FDD), (Release 1999)," 3GPP, Technical Specification Group Radio Access Network, ETSI TS 25212 V3.5.0, 2000.



Yiqing Zhou (M'04) was born in Zhejiang, China, in 1975. She received the B.S. degree in communication and information engineering and the M.S. degree in signal and information processing from Southeast University, Nanjing, China, in 1997 and 2000, respectively, and the Ph.D. degree in electrical and electronic engineering from the University of Hong Kong, Hong Kong, in February 2004.

Since June 2004, she has been with the Department of Electrical and Electronic Engineering, University of Hong Kong as a Postdoctoral Fellow. Her research interests include coding theory, spread spectrum, OFDM systems, interference cancellation, hybrid ARQ, and other transmission techniques for wireless high-speed data communications.



Jiangzhou Wang (M'91–SM'94) received the B.S. and M.S. degrees from Xidian University, Xian, China, in 1983 and 1985, respectively, and the Ph.D. degree (with Greatest Distinction) from the University of Ghent, Ghent, Belgium, in 1990, all in electrical engineering.

He is currently a Professor and Chair in the Department of Electronics, University of Kent, Kent, U.K. From 1995 to 2005, he was with the University of Hong Kong, where he is still serving as an Honorary Professor. From 1992 to 1995, he

was a Senior System Engineer at Rockwell International Corporation (now Conexant), CA. From 1990 to 1992, he was a Postdoctoral Fellow in the University of California, San Diego. He has held a Visiting Professor position in NTT DoCoMo, Japan. He has written/edited two books, entitled *Broadband Wireless Communications* (Boston, MA: Kluwer, 2001) and *Advances in 3G Enhanced Technologies for Wireless Communications* (Norwood, MA: Artech House, 2002). He holds one U.S. patent in the GSM system. He has published over 140 papers, including more than 40 IEEE Transactions/Journal papers in the areas of wireless mobile and spread-spectrum communications.

Dr. Wang was a Technical Chairman of the IEEE Workshop in 3G Mobile Communications in 2000. He is an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Wideband CDMA, 2000 and 2001, and Advances in Multicarrier CDMA, 2006). He is listed in *Who's Who in the World*.