

MIMO-OFCDM Systems with Joint Iterative Detection and Optimal Power Allocation

Yiqing Zhou, *Member, IEEE*, and Tung-Sang Ng, *Fellow, IEEE*

Abstract—This paper investigates the orthogonal frequency and code division multiplexing (OFCDM) systems with multiple input multiple output multiplexing (MIMO-OFCDM) and multi-code transmission. Combining the iterative detection in the space domain and the hybrid multi-code interference (MDI) cancellation and minimum mean square error (MMSE) detection in the frequency domain, a joint iterative detection is proposed, which enables space and frequency diversity gains to be jointly exploited. Moreover, using a two-dimensional (2-D) averaging channel estimation algorithm, a close form expression is derived for the optimal power allocation between the pilot and all data channels that achieves the best system performance. It is shown that the optimal power ratio mainly depends on the channel estimation algorithm, the number of transmit antennas as well as the number of pilot and data symbols in a packet, but is not sensitive to the changes in signal-to-noise ratio (SNR) and diversity gains. Simulations are conducted to verify the derived optimal power ratio and study the performance of the proposed joint detection algorithm. It is shown that considerable improvement can be obtained when the number of loops in the joint iterative detection increases. Moreover, the system performance is enhanced significantly when the frequency domain spreading factor, N_F , increases.

Index Terms—OFCDM, MIMO, optimal power allocation, two-dimensional spreading, multi-code, iterative interference cancellation, channel estimation.

I. INTRODUCTION

MULTIPLE-INPUT and multiple-output (MIMO) multiplexing is an efficient technique to increase the data rate. By transmitting independent data on different antennas simultaneously, MIMO can increase the data rate linearly with the minimum number of transmit and receive antennas [1]–[2]. On the other hand, combining OFDM with two-dimensional (2-D) spreading [3]–[4], orthogonal frequency and code division multiplexing (OFCDM) [5] is a promising wireless access technique for the broadband downlink transmission. Based on OFDM, OFCDM can combat the severe multipath interference. Moreover, 2-D spreading provides additional benefits. With the introduction of time domain spreading, the system can provide flexible transmission rates. Furthermore, frequency diversity gain is achieved through frequency domain spreading. Although frequency diversity gain can be obtained by channel coding as well, the gain is limited due to the

limited coding rate and constraint length [6]–[7]. Therefore, even in channel coded systems, frequency domain spreading should still be employed to provide additional diversity gain. Meanwhile, because of 2-D spreading, the data rate of the OFCDM system is reduced to $1/N$, where $N = N_T \times N_F$ is the total spreading factor, and N_T and N_F are the time domain and frequency domain spreading factors, respectively. To increase the data rate, multi-code transmission can be employed. Using orthogonal variable spreading factor (OVSF) codes [8] as the spreading codes, there are N_T and N_F different codes in the time and frequency domains, respectively. Hence, up to $N = N_T \times N_F$ orthogonal 2-D codes are available, and the system can provide various service rates by assigning different number of 2-D codes to one single user.

The OFCDM system with single antenna (SISO-OFCDM) has been widely investigated [9]–[11]. However, there are only few published works on the MIMO-OFCDM system. With only time domain spreading, the performance of MIMO-OFCDM was investigated in [12]–[13], where a sub-optimum maximum likelihood detection (MLD) method was considered. In [14], to resolve the near-far problem, transmission power was optimized for a convolutionally coded MIMO-OFCDM system using single-code transmission. This paper considers the MIMO-OFCDM system with 2-D spreading and multi-code transmission, where the desired data signals are interfered by not only co-channel signals from other transmit antennas (or multi-antenna interference (MTI)), but also multi-code interference (MDI) from other code channels. Although the MLD achieves the optimal performance, it is practically infeasible since its complexity increases exponentially with the number of antennas and code channels. On the other hand, various sub-optimum detection algorithms have been proposed for MIMO systems, such as successive interference cancellation (SIC) [15]–[17]. Based on zero-forcing SIC (ZF-SIC), a multi-loop iterative detection in the space domain was proposed [18]. Moreover, in the frequency domain, a multi-stage hybrid MDI cancellation and MMSE detection provides superior performance to the pure MMSE (or conventional MMSE) detection [10]–[11]. Combining the iterative detection in the space domain and the hybrid detection in the frequency domain, a joint iterative signal detection is proposed. It will be shown that the performance of the MIMO-OFCDM system is significantly enhanced due to the frequency and space diversity gains obtained by using the joint detection.

Meanwhile, in SISO-OFCDM, it has been shown by simulation and numerical evaluation that the quality of channel estimation is critical to the system performance [10]–[11]. It is therefore important to investigate the performance of

Manuscript received December 29, 2007; revised April 29, 2008 and July 16, 2008; accepted July 17, 2008. The associate editor coordinating the review of this paper and approving it for publication was M. Morelli.

The authors are with the Department of Electrical and Electronic Engineering, University of Hong Kong, Pokfulam Road, Hong Kong (e-mail: {yqzhou, tsng}@eee.hku.hk).

This work was supported by a grant HKU7164/04E from The Research Grants Council of Hong Kong.

Digital Object Identifier 10.1109/T-WC.2008.071462

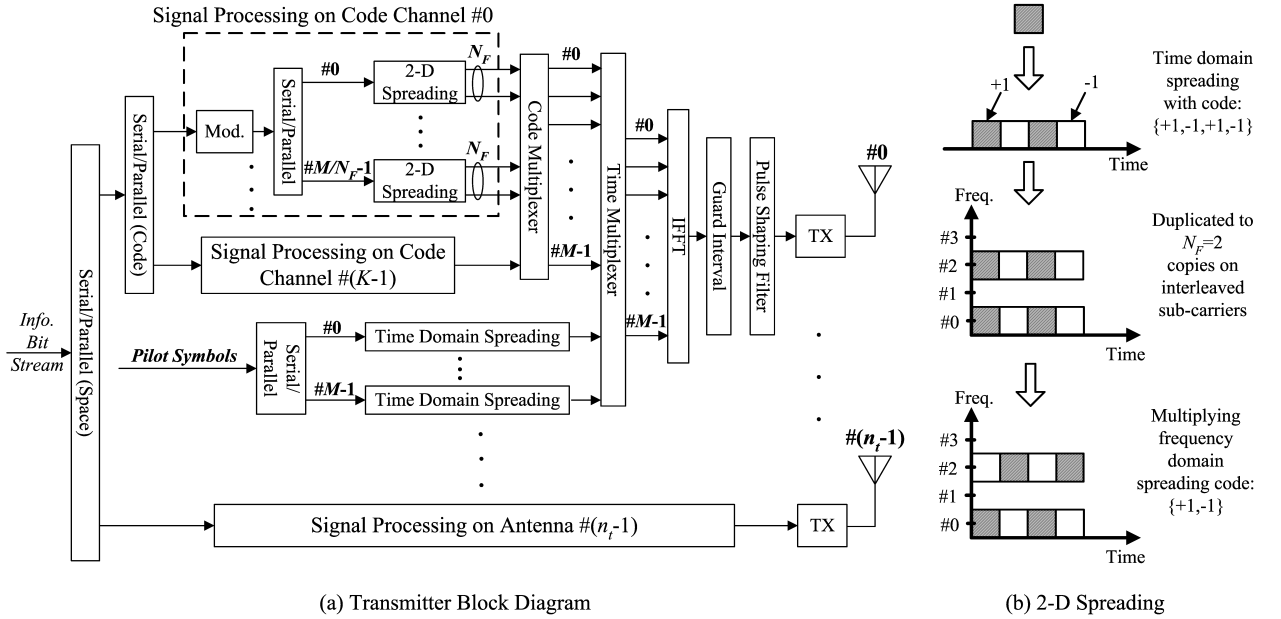


Fig. 1. Transmitter structure of the MIMO-OFCDM system

MIMO-OFCDM systems with imperfect channel estimation. Under limited transmission power at each antenna, the channel estimation is optimized by using the optimal power allocation (or power ratio) between the pilot and all data channels (β/K) that achieves the best system performance. By minimizing the mean square error (MSE) during signal detection, a close form expression for the optimal power ratio β/K is derived in this paper. The optimal power ratio is shown to depend on the channel estimation algorithm, the number of transmit antennas as well as the number of pilot and data symbols in a packet, but is not sensitive to other system parameters such as signal-to-noise ratio (SNR), the number of receive antennas, and frequency domain spreading factor.

The rest of the paper is organized as follows. Section II introduces the basic structure of the MIMO-OFCDM system and the channel model. Section III describes the channel estimation algorithm and the joint iterative detection algorithm. A close form expression is derived in Section IV for the optimal power ratio between the pilot channel and all data channels. Then simulation results are presented in Section V to verify the derived optimal power ratio and demonstrate the performance of MIMO-OFCDM with the joint detection. Finally, conclusions are drawn in the last section.

II. SYSTEM DESCRIPTION

A. System Model

The transmitter of the MIMO-OFCDM system is shown in Fig.1(a). Information bit streams are firstly serial to parallel (S/P) converted into n_t sub-streams, corresponding to n_t transmit antennas. At each antenna, information bits are processed by a 2-D spread OFCDM with multi-code transmission, where each sub-stream is further S/P converted into K streams, corresponding to K data codes. For each code, data bits are mapped to symbols and S/P converted into M/N_F symbol streams, where M is the total number of sub-carriers. Using the dedicated code, each data symbol is 2-D spread with

a factor of $N = N_T \times N_F$. Assuming $N = 4 \times 2$, an example of 2-D spreading is shown in Fig.1(b). The data symbol is firstly spread in the time domain, then duplicated into N_F copies in the frequency domain and multiplied by the frequency domain spreading code. In order to achieve good frequency diversity gain, N_F interleaved sub-carriers are employed instead of N_F consecutive sub-carriers. After 2-D spreading, a code multiplexer is employed to combine the signals spread by K different 2-D codes, realizing the multi-code transmission. To provide high data rate, a full loaded system with $K = N$ is considered. At the same time, a time-multiplexed pilot channel [12] is employed for channel estimation. Known QPSK modulated pilot symbols are S/P converted into M sub-streams and spread in the time domain using OVFS codes with length N_P . The spread pilot signals pass through the time multiplexer with data signals, followed by the IFFT block, guard interval insertion and the pulse shaping filter. It should be noted that other pilot structures can also be employed. For example, when different antennas transmit pilot symbols over distinct sub-carriers, orthogonal pilot channels among transmit antennas are realized in the frequency domain within one single OFCDM symbol duration. Then, channel information on each sub-carrier can be obtained from the ML estimator or MMSE estimator [19]. Since the proposed approach of deriving the optimal power ratio in Sec. IV can be applied to different pilot structures and channel estimation algorithms, the time multiplexed pilot structure is employed as an example.

The resultant signal packet structure at each antenna is shown in Fig.2. The packet structure is the same at all transmit antennas, except that the pilot channels at different antennas are spread with different OVFS codes, so that they can keep orthogonal to each other in the code domain after passing through a slow fading channel. In the time domain, the packet starts with N_P OFCDM symbols for pilot, followed by N_D symbols for data transmission. The duration of each OFCDM

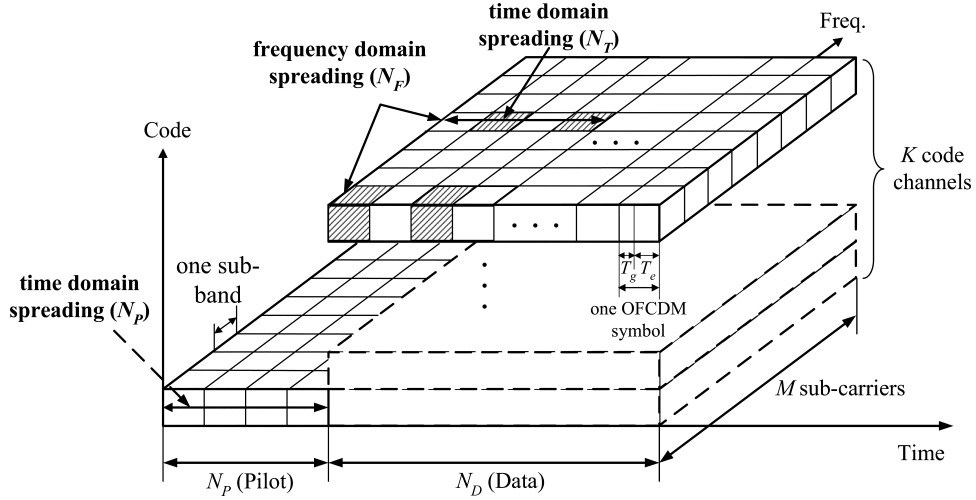


Fig. 2. Packet structure at each antenna

$$S_p(t) = \begin{cases} \sum_{m=0}^{M-1} d_{pilot,p,m} c_{N_P,i}^{(p)} e^{j2\pi m(t-T_g-iT_s)/T_e} f(t-T_g-iT_s) & i \in [0, N_P - 1] \\ \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} d_{p,i,m,k} C_{i,m,k} e^{j2\pi m(t-T_g-iT_s)/T_e} f(t-T_g-iT_s) & i \in [N_P, N_P + N_D - 1] \end{cases}, \quad (1)$$

symbol is $T_s = T_e + T_g$, with T_e and T_g denoting the effective OFCDM symbol duration and guard interval, respectively. In the frequency domain, there are totally M sub-carriers and pilot symbols are adopted at each sub-carrier. Finally, in the code domain, only one time domain spreading code is used by the pilot channel, while K 2-D codes are employed for data transmission.

Accordingly, the transmitted signal at the p^{th} ($0 \leq p \leq n_t - 1$) antenna in one packet duration is given by (1), where $d_{pilot,p,m}$ is the known pilot symbol on the m^{th} sub-carrier in the i^{th} OFCDM symbol duration, $c_{N_P,i}^{(p)}$ is the i^{th} chip of the p^{th} OVFSF code with length N_P on the code tree [8], $d_{p,i,m,k}$ is the data symbol of the k^{th} code channel, $C_{i,m,k}$ stands for the 2-D spreading code chip, and $f(t)$ is the rectangular pulse shaping filter. The powers of one data symbol and one pilot symbol are set to $E\{|d_{p,i,m,k}|^2\} = P_d$ and $E\{|d_{pilot,p,m}|^2\} = \beta P_d$, respectively. Assume that a block interleaver with depth N_F is used in the frequency domain. It follows that the 0^{th} , $(M/N_F)^{\text{th}}$, \dots , and $(M - M/N_F)^{\text{th}}$ sub-carrier carry the same data symbol, the 1^{st} , $(M/N_F + 1)^{\text{th}}$, \dots , and $(M - M/N_F + 1)^{\text{th}}$ sub-carrier carry another data symbol, and so on. Meanwhile, in the time domain, the N_P^{th} , $(N_P + 1)^{\text{th}}$, \dots , and $(N_P + N_T - 1)^{\text{th}}$ OFCDM symbols carry the same data symbol, while the $(N_P + N_T)^{\text{th}}$, \dots , and $(N_P + 2N_T - 1)^{\text{th}}$ OFCDM symbols carry another data symbol, and so on.

B. Channel Model

Although the broadband MIMO-OFCDM system experiences a highly frequency selective channel, the signal transmitted on each sub-carrier experiences a flat fading channel. Assuming a slow fading channel which is approximately fixed

in one packet duration, $h_{q,p,m}$ denotes the complex channel fading on the m^{th} sub-carrier from the p^{th} transmit antenna to the q^{th} receive antenna. The amplitude and phase of $h_{q,p,m}$ are assumed to be Rayleigh distributed with $P_{ch} = E\{|h_{q,p,m}|^2\}$ and uniformly distributed in $[0, 2\pi)$, respectively. Furthermore, the sub-carrier correlation coefficient in the frequency domain between h_{q,p,m_0} and h_{q,p,m_1} is given by [20]

$$\begin{aligned} \rho_{m_0,m_1} &= \frac{E\{h_{q,p,m_0}^* h_{q,p,m_1}\}}{\sqrt{E\{|h_{q,p,m_0}|^2\}} \sqrt{E\{|h_{q,p,m_1}|^2\}}} \\ &= \frac{1}{\sqrt{1 + (\Delta f / \Delta f_c)^2}}, \end{aligned} \quad (2)$$

where $(\cdot)^*$ stands for the conjugate operation, Δf is the frequency separation between the m_0^{th} and m_1^{th} sub-carriers and Δf_c is the coherence bandwidth of the channel. For different transmit or receive antennas, the channel fading is assumed to be independent to each other.

III. JOINT ITERATIVE SIGNAL DETECTION

A. Channel Estimation

At the receiver, after FFT, the pilot signal ($\bar{i} \in [0, N_P - 1]$) on the \bar{m}^{th} sub-carrier of the \bar{q}^{th} antenna is given by

$$y_{\bar{q},\bar{i},\bar{m}} = \sum_{p=0}^{n_t-1} h_{\bar{q},p,\bar{m}} d_{pilot,p,\bar{m}} c_{N_P,\bar{i}}^{(p)} + \eta_{\bar{q},\bar{i},\bar{m}} \quad (3)$$

while the data signal ($\bar{i} \in [N_P, N_P + N_D - 1]$) is given by

$$y_{\bar{q},\bar{i},\bar{m}} = \sum_{p=0}^{n_t-1} h_{\bar{q},p,\bar{m}} \sum_{k=0}^{K-1} d_{p,\bar{i},\bar{m},k} C_{\bar{i},\bar{m},k} + \eta_{\bar{q},\bar{i},\bar{m}} \quad (4)$$

$$\begin{aligned}\bar{h}_{\bar{q},\bar{p},\bar{m}} &= \frac{1}{d_{\text{pilot},\bar{p},\bar{m}}} \left(\frac{1}{N_P} \sum_{\bar{i}=0}^{N_P-1} y_{\bar{q},\bar{i},\bar{m}} c_{N_P,\bar{i}}^{(\bar{p})} \right) \\ &= \frac{1}{d_{\text{pilot},\bar{p},\bar{m}}} \left(\sum_{p=0}^{n_t-1} h_{\bar{q},p,\bar{m}} d_{\text{pilot},p,\bar{m}} \left(\frac{1}{N_P} \sum_{\bar{i}=0}^{N_P-1} c_{N_P,\bar{i}}^{(p)} c_{N_P,\bar{i}}^{(\bar{p})} \right) + \frac{1}{N_P} \sum_{\bar{i}=0}^{N_P-1} \eta_{\bar{q},\bar{i},\bar{m}} c_{N_P,\bar{i}}^{(\bar{p})} \right) = h_{\bar{q},\bar{p},\bar{m}} + \nu_{\bar{q},\bar{p},\bar{m}}\end{aligned}\quad (5)$$

$$\begin{aligned}\text{MSE}_h &= E \left\{ \left| \tilde{h}_{\bar{q},\bar{p},\bar{m}} - h_{\bar{q},\bar{p},\bar{m}} \right|^2 \right\} = \left\{ \left| \sum_{i=-W}^W \tilde{\alpha}_i h_{\bar{q},\bar{p},\bar{m}+i} + \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \sum_{i=-W}^W \alpha_i \nu_{\bar{q},\bar{p},\bar{m}+i} \right|^2 \right\} \\ &= P_{ch} \sum_{i_0=-W}^W \sum_{i_1=-W}^W \tilde{\alpha}_{i_0} \tilde{\alpha}_{i_1}^* \rho_{i_0,i_1} + \sigma_H^2 \cdot \left| \sum_{i=-W}^W \alpha_i \right|^{-2} \sum_{i=-W}^W |\alpha_i|^2 = P_{ch} B_H + \sigma_H^2 A_H\end{aligned}\quad (8)$$

$$\begin{aligned}z_{q,\bar{m}} &= \frac{1}{N_T} \sum_{\bar{i}=N_P}^{N_P+N_T-1} y_{q,\bar{i},\bar{m}} c_{N_T,(\bar{i}-N_P)\%N_T}^{(0)} \\ &= h_{q,\bar{p},\bar{m}} \left(\underbrace{d_{\bar{p},\bar{k}} c_{N_F,\bar{m}}^{(\bar{k})}}_{MTI} \right) + \underbrace{\sum_{\substack{p=0 \\ p \neq \bar{p}}}^{n_t-1} h_{q,p,\bar{m}} \left(\sum_{k=0}^{K_C} d_{p,k} c_{N_F,\bar{m}}^{(k)} \right)}_{MDI} + h_{q,\bar{p},\bar{m}} \left(\sum_{\substack{k=0 \\ k \neq \bar{k}}}^{K_C} d_{\bar{p},k} c_{N_F,\bar{m}}^{(k)} \right) + \varepsilon_{q,\bar{m}}\end{aligned}\quad (9)$$

where $\eta_{\bar{q},\bar{i},\bar{m}}$ is the background noise term with zero mean and variance σ_n^2 . To obtain precise channel information, a two-dimensional (2-D: time and frequency domains) average channel estimation is employed [21]. Firstly, using the time domain despreading, a preliminary estimation of the channel on the \bar{m}^{th} sub-carrier from the \bar{p}^{th} transmit antenna to the \bar{q}^{th} receive antenna is given by (5), where $\nu_{\bar{q},\bar{p},\bar{m}}$ is the AWGN noise with zero mean and variance

$$\begin{aligned}\sigma_H^2 &= E \left\{ |\nu_{\bar{q},\bar{p},\bar{m}}|^2 \right\} = \frac{1}{|d_{\text{pilot},\bar{p},\bar{m}}|^2 N_P^2} \cdot \\ &\quad \sum_{i_0=0}^{N_P-1} \sum_{i_1=0}^{N_P-1} c_{N_P,i_0}^{(\bar{p})} c_{N_P,i_1}^{(\bar{p})} E \left\{ \eta_{q,i_0,\bar{m}} \eta_{q,i_1,\bar{m}}^* \right\} \\ &= \sigma_n^2 / (\beta P_d N_P)\end{aligned}\quad (6)$$

where $|d_{\text{pilot},\bar{p},\bar{m}}|^2 = \beta P_d$ because of QPSK modulated pilot symbols. The preliminary estimation $\bar{h}_{\bar{q},\bar{p},\bar{m}}$ can be further smoothed in the frequency domain by exploiting the correlation of channel fading among adjacent sub-carriers

$$\begin{aligned}\tilde{h}_{\bar{q},\bar{p},\bar{m}} &= \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \sum_{i=-W}^W \alpha_i \bar{h}_{\bar{q},\bar{p},\bar{m}+i} \\ &= \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \sum_{i=-W}^W \alpha_i h_{\bar{q},\bar{p},\bar{m}+i} \\ &\quad + \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \sum_{i=-W}^W \alpha_i \nu_{\bar{q},\bar{p},\bar{m}+i}\end{aligned}\quad (7)$$

where $\alpha_0 = 1$ and $\alpha_i = \alpha_{-i}$ for $i = 1, \dots, W$. The performance of this 2-D average channel estimation algorithm has been investigated in [21]. It was shown that in highly frequency selective channels, a proper size for the frequency domain averaging should be $W = 1$ and α_1 should take a

value between 0.6 and 1.0. Letting $\tilde{\alpha}_i = \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \alpha_i$ for $i = \pm 1, \dots, \pm W$ and $\tilde{\alpha}_0 = \left(\sum_{i=-W}^W \alpha_i \right)^{-1} \alpha_0 - 1$, the MSE of the channel estimation is given by (8), where $P_{ch} = E \left\{ |h_{q,p,m}|^2 \right\}$, $B_H = \sum_{i_0=-W}^W \sum_{i_1=-W}^W \tilde{\alpha}_{i_0} \tilde{\alpha}_{i_1}^* \rho_{i_0,i_1}$ and $A_H = \left| \sum_{i=-W}^W \alpha_i \right|^{-2} \sum_{i=-W}^W |\alpha_i|^2$.

B. Joint Iterative Detection

Without loss of generality, consider the data signals received on the 0^{th} , $(M/N_F)^{\text{th}}$, \dots , and $(M - M/N_F)^{\text{th}}$ sub-carriers from the N_P^{th} to the $(N_P + N_T - 1)^{\text{th}}$ OFCDM symbol duration, carrying the same unspread data symbol $d_{\bar{p},0,0,\bar{k}}$ (denoted as $d_{\bar{p},\bar{k}}$ hereafter). The desired data is assumed to be spread by a 2-D code of $\{C_{N_T}^{(0)}, C_{N_F}^{(\bar{k})}\}$. Then, the rest $(K - 1)$ 2-D codes are divided into two subsets: one set including 2-D codes with the same time domain spreading code $C_{N_T}^{(0)}$ but different frequency domain spreading codes, i.e., $\Omega_F = \left\{ \{C_{N_T}^{(0)}, C_{N_F}^{(k_F)}\} \mid k_F \neq \bar{k} \right\}$, and the other set including 2-D codes with different time domain spreading codes, i.e., $\Omega_T = \left\{ \{C_{N_T}^{(k_T)}, C_{N_F}^{(k_F)}\} \mid k_T \neq 0 \right\}$ [10]. 2-D codes in different sub-sets (Ω_F or Ω_T) have different contribution to the MDI to the desired code $\{C_{N_T}^{(0)}, C_{N_F}^{(\bar{k})}\}$. In highly frequency selective channels, codes from Ω_F cause severe MDI to the desired code because their orthogonality in the frequency domain is distorted by the frequency selective fading on sub-carriers. On the other hand, in a slow fading channel, the orthogonality in the time domain between any

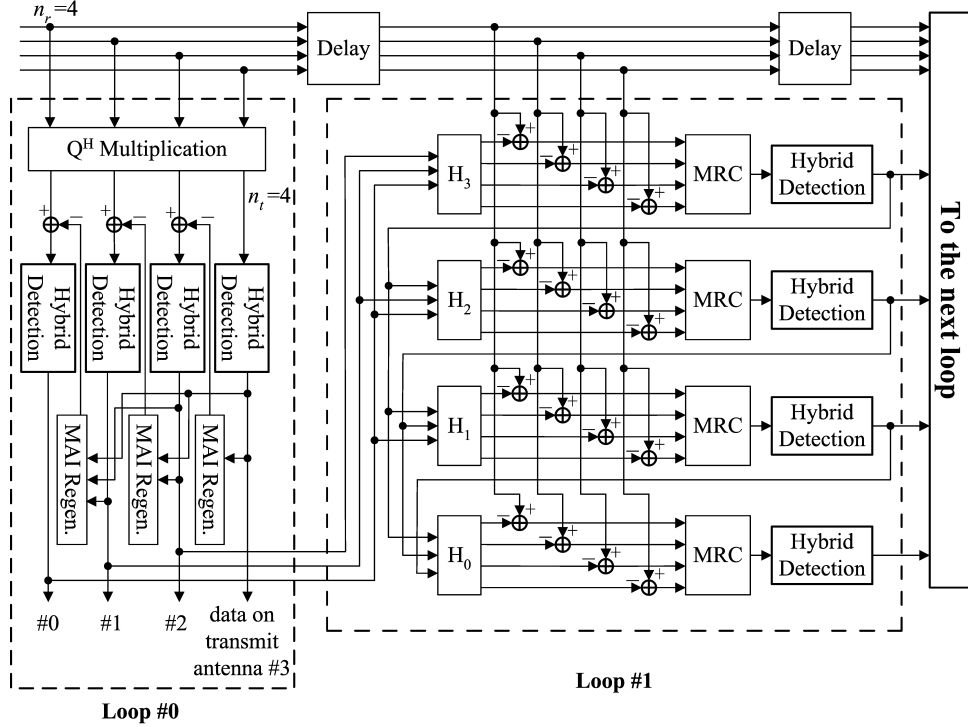


Fig. 3. Example of joint space-frequency detection for $n_t = n_r = 4$

codes from Ω_T and the desired code is maintained. Thus, there is no MDI from Ω_T . To keep the MDI small, the 2-D codes with different time domain spreading codes should be assigned for data transmission first [22]. Using this code partition, the signal after time domain despreading is given by (9), where K_C ($K_C \leq N_F - 1$) is the number of codes in Ω_F , and $\varepsilon_{q,\bar{m}}$ is the noise term with zero mean and variance σ_n^2/N_T . It can be seen that the desired signal $d_{\bar{p},k}$ is corrupted by MTI in the space domain generated by signals sending by other transmit antennas, as well as MDI in the code domain generated by signals radiated by the same antenna but on different 2-D codes from Ω_F .

A joint iterative signal detection algorithm is proposed for the MIMO-OFCDM system to recover the desired data from the outputs of the time domain despreader. An example with four transmit and receive antennas is shown in Fig.3. It can be seen that the algorithm is composed of multi-loop iterative detection in the space domain with multi-stage hybrid detection in the frequency domain within each loop. In the 0th loop, ZF-SIC is employed in the space domain. At each antenna, hybrid detection is used to further detect the data symbol transmitted on each code. In later loops, since tentative data decisions have been obtained in the previous loop, a MTI cancellation can be carried out, after which the multi-code signal at each transmit antenna can be collected by maximum ratio combining (MRC). A detailed description of the joint iterative detection is given as follows.

1) *Operations in the 0th Loop:* Define $\bar{Z}_{\bar{m}} = (z_{0,\bar{m}}, \dots, z_{n_r-1,\bar{m}})^T$ as the time domain despreader outputs at the n_r receive antennas on the \bar{m}^{th} sub-carrier. $\bar{D}_{\bar{m}} = (\sum_{k=0}^{K_C} d_{0,k} c_{N_F,\bar{m}}^{(k)}, \dots, \sum_{k=0}^{K_C} d_{n_t-1,k} c_{N_F,\bar{m}}^{(k)})^T$ stands for

the $n_t \times 1$ transmitted signal column vector. Hence, the time domain despreader outputs can be rewritten in matrix form as $\bar{Z}_{\bar{m}} = \bar{H}_{\bar{m}} \bar{D}_{\bar{m}} + \bar{N}_{\bar{m}}$, where $\bar{H}_{\bar{m}} = (\bar{h}_{\bar{m},0}, \dots, \bar{h}_{\bar{m},n_t-1})$ represents the $n_r \times n_t$ channel matrix on the \bar{m}^{th} sub-carrier, $\bar{h}_{\bar{m},p} = (h_{0,p,\bar{m}}, \dots, h_{n_r-1,p,\bar{m}})^T$ is a $n_r \times 1$ channel vector, and $\bar{N}_{\bar{m}} = (\varepsilon_{0,\bar{m}}, \varepsilon_{1,\bar{m}}, \dots, \varepsilon_{n_r-1,\bar{m}})^T$ is a noise vector. In the 0th loop of the joint iterative detection, ZF-SIC is employed in the space domain. The QR decomposition of the estimated channel matrix $\hat{H}_{\bar{m}}$ is expressed as $\hat{H}_{\bar{m}} = \hat{Q}_{\bar{m}} \hat{R}_{\bar{m}}$, where the element of $\hat{H}_{\bar{m}}$ is given by (7), $\hat{Q}_{\bar{m}}$ is a $n_r \times n_t$ orthogonal matrix and $\hat{R}_{\bar{m}}$ is a $n_t \times n_t$ upper triangular matrix. After multiplying $\bar{Z}_{\bar{m}}$ with the hermitian transpose of $\hat{Q}_{\bar{m}}$, i.e., $\hat{Q}_{\bar{m}}^H$, the signal is given by

$$\begin{aligned} \bar{Y}_{\bar{m}} &= \hat{Q}_{\bar{m}}^H \bar{Z}_{\bar{m}} = \hat{Q}_{\bar{m}}^H (\hat{H}_{\bar{m}} + V_{\bar{m}}) \bar{D}_{\bar{m}} + \hat{Q}_{\bar{m}}^H \bar{N}_{\bar{m}} \\ &= \hat{R}_{\bar{m}} \bar{D}_{\bar{m}} + \Delta_{\bar{m}}, \end{aligned} \quad (10)$$

where $V_{\bar{m}} = \bar{H}_{\bar{m}} - \hat{H}_{\bar{m}}$ stands for the channel estimation error and $\Delta_{\bar{m}} = \hat{Q}_{\bar{m}}^H V_{\bar{m}} \bar{D}_{\bar{m}} + \hat{Q}_{\bar{m}}^H \bar{N}_{\bar{m}}$ is the total noise vector. The covariance matrix of $\Delta_{\bar{m}}$ can be obtained as

$$\begin{aligned} E \{ \Delta_{\bar{m}} \Delta_{\bar{m}}^H \} &= \left[(K_C + 1) P_d n_t \cdot MSE_h + \frac{\sigma_n^2}{N_T} \right] I_{n_t} \\ &= \sigma^2 I_{n_t} \end{aligned} \quad (11)$$

where MSE_h is the MSE of the channel estimation given by (8) and I_{n_t} is the $n_t \times n_t$ identity matrix. Due to the upper triangular structure of $\hat{R}_{\bar{m}}$, the \bar{p}^{th} element of $\bar{Y}_{\bar{m}}$ can be expressed as

$$\gamma_{\bar{p},\bar{m}} = \hat{r}_{\bar{m},\bar{p},\bar{p}} \sum_{k=0}^{K_C} d_{\bar{p},k} c_{N_F,\bar{m}}^{(k)} + \delta_{\bar{p},\bar{m}}$$

$$\begin{aligned}
\gamma_{n_t-1,\bar{m},\bar{k}}^{(s)} &= \gamma_{n_t-1,\bar{m}} - \hat{r}_{\bar{m},n_t-1,n_t-1} \sum_{k=0,k \neq \bar{k}}^{K_C} \hat{d}_{n_t-1,k,s-1} c_{N_F,\bar{m}}^{(k)} \\
&= \hat{r}_{\bar{m},n_t-1,n_t-1} d_{n_t-1,\bar{k}} c_{N_F,\bar{m}}^{(k)} + \underbrace{\hat{r}_{\bar{m},n_t-1,n_t-1} \sum_{k=0,k \neq \bar{k}}^{K_C} (d_{n_t-1,k} - \hat{d}_{n_t-1,k,s-1}) c_{N_F,\bar{m}}^{(k)}}_{\text{residual MDI}} + \delta_{n_t-1,\bar{m}}
\end{aligned} \quad (16)$$

$$l_{n_t-1,\bar{m},\bar{k}}^{(s)} = \frac{E \left\{ d_{n_t-1,\bar{k}} \left(\gamma_{n_t-1,\bar{m},\bar{k}}^{(s)} \right)^* \right\}}{E \left\{ \left| \gamma_{n_t-1,\bar{m},\bar{k}}^{(s)} \right|^2 \right\}} = \frac{P_d \cdot \hat{r}_{\bar{m},n_t-1,n_t-1}^* c_{N_F,\bar{m}}^{(\bar{k})}}{(K_C + 1)^{-1} (N_D/N_T)^{-1} \sum_{k=0}^{K_C} \sum_{j=0}^{(N_D/N_T)-1} \left| \gamma_{n_t-1,\bar{m},\bar{k}}^{(s)}(j) \right|^2} = \omega_{n_t-1,\bar{m}}^{(s)} c_{N_F,\bar{m}}^{(\bar{k})} \quad (17)$$

$$\begin{aligned}
y_{n_t-1,\bar{k}}^{(s)} &= \sum_{\bar{m}=0,(M/N_F),\dots,(M-M/N_F)} l_{n_t-1,\bar{m},\bar{k}}^{(s)} \cdot \gamma_{n_t-1,\bar{m},\bar{k}}^{(s)} = \left(\sum_{\bar{m}} \omega_{n_t-1,\bar{m}}^{(s)} \hat{r}_{\bar{m},n_t-1,n_t-1} \right) d_{n_t-1,\bar{k}} \\
&\quad + \underbrace{\sum_{\bar{m}} \omega_{n_t-1,\bar{m}}^{(s)} c_{N_F,\bar{m}}^{(\bar{k})} \varepsilon_{n_t-1,\bar{m}}}_{\text{Noise}} + \underbrace{\sum_{\bar{m}} \sum_{k=0,k \neq \bar{k}}^{K_C} \omega_{n_t-1,\bar{m}}^{(s)} c_{N_F,\bar{m}}^{(\bar{k})} \hat{r}_{\bar{m},n_t-1,n_t-1} (d_{n_t-1,k} - \hat{d}_{n_t-1,k,s-1}) c_{N_F,\bar{m}}^{(k)}}_{\text{residual MDI}}
\end{aligned} \quad (18)$$

$$+ \sum_{p=\bar{p}+1}^{n_t-1} \left(\hat{r}_{\bar{m},p,\bar{p}} \sum_{k=0}^{K_C} d_{p,k} c_{N_F,\bar{m}}^{(k)} \right) \quad (12)$$

where $\hat{r}_{\bar{m},i,j}$ is the element on the i^{th} row and j^{th} column of $\hat{R}_{\bar{m}}$, and $\delta_{\bar{p},\bar{m}}$ is the \bar{p}^{th} element of the noise vector with zero mean and variance σ^2 . It can be seen that for $\bar{p} = n_t - 1$, $\gamma_{n_t-1,\bar{m}}$ is free of MTI and can be used to estimate the signals at the $(n_t - 1)^{\text{th}}$ transmit antenna. Let $\vec{d}_{n_t-1,F} = (d_{n_t-1,0}, \dots, d_{n_t-1,K_C})^T$ and $\vec{N}_{\delta} = (\delta_{n_t-1,0}, \dots, \delta_{n_t-1,M-M/N_F})^T$ denote the data symbols carried by the $K_C + 1$ codes and the channel noise on the N_F interleaved sub-carriers, respectively, $H_F = \text{diag}(\hat{r}_{0,n_t-1,n_t-1}, \dots, \hat{r}_{(M-M/N_F),n_t-1,n_t-1})$ be a $N_F \times N_F$ diagonal matrix, and $C_F = ((c_{0,0}, \dots, c_{0,N_F-1})^T, \dots, (c_{K_C,0}, \dots, c_{K_C,N_F-1})^T)^T$ represent the $N_F \times (K_C + 1)$ OVFSF code matrix. The signals on the N_F interleaved sub-carriers at $\bar{p} = n_t - 1$, i.e., $\Gamma_{n_t-1} = (\gamma_{n_t-1,0}, \dots, \gamma_{n_t-1,M-(M/N_F)})^T$, can then be expressed as

$$\Gamma_{n_t-1} = H_F C_F \vec{d}_{n_t-1,F} + \vec{N}_{\delta} \quad (13)$$

A multi-stage hybrid MDI cancellation and MMSE detection [10] is employed in the frequency domain to recover the data symbols transmitted by the $(n_t - 1)^{\text{th}}$ antenna. At the 0^{th} stage of hybrid detection, the input signal is processed directly by a pure MMSE combining without MDI cancellation, whose weighting matrix is given by

$$L_{MMSE} = P_d (H_F C_F)^H \left(P_d H_F C_F (H_F C_F)^H + \sigma^2 I \right)^{-1} \quad (14)$$

Assuming a full-loaded MIMO-OFCDM system with $K = N$, (14) can be simplified and the MMSE weight on the \bar{m}^{th} sub-

carrier for the \bar{k}^{th} code is given by

$$\begin{aligned}
l_{n_t-1,\bar{m},\bar{k}}^{(0)} &= \frac{P_d \cdot \hat{r}_{\bar{m},n_t-1,n_t-1}^* c_{N_F,\bar{m}}^{(\bar{k})}}{N_F P_d \cdot \left| \hat{r}_{\bar{m},n_t-1,n_t-1} \right|^2 + \sigma^2} \\
&= \omega_{n_t-1,\bar{m}}^{(0)} c_{N_F,\bar{m}}^{(\bar{k})}
\end{aligned} \quad (15)$$

After the MMSE detection, tentative decisions obtained at the 0^{th} stage are fed to a MDI regenerator. Then, at the 1^{st} stage of the hybrid detection, the regenerated MDI is subtracted from the input signal. The resultant signal is further processed by the MMSE, whose weights are updated stage by stage due to the reduction of MDI at each stage. Similarly, new data decisions are obtained at the 1^{st} stage and are used to regenerate the MDI for MDI cancellation in the next stage. Denote $\hat{d}_{n_t-1,k,s-1}$ as the tentative data decisions obtained at the $(s - 1)^{\text{th}}$ stage of hybrid detection, which is zero for $s = 0$. Then, after the MDI cancellation at the s^{th} stage, the signal on the \bar{k}^{th} code is given by (16). Moreover, on the \bar{m}^{th} sub-carrier, the MMSE weight at the s^{th} stage ($s \geq 0$) can be expressed as (17). Note that the ensemble average $E \left\{ \left| \gamma_{n_t-1,\bar{m},\bar{k}}^{(s)} \right|^2 \right\}$ is estimated with time and code average [10]-[11], where $\gamma_{n_t-1,\bar{m},\bar{k}}^{(s)}(j)$ stands for the j^{th} unspread data signal in the time domain in one packet duration. Thus, the decision variable for the data on the \bar{k}^{th} code after the MMSE combining at the s^{th} stage is given by (18). Given $y_{n_t-1,\bar{k}}^{(s)}$, the data decisions at the s^{th} stage can be obtained as $\hat{d}_{n_t-1,k,s}$. When the multi-stage hybrid detection is completed, the data transmitted on the $(n_t - 1)^{\text{th}}$ antenna are recovered as $\hat{d}_{n_t-1,0}^{(0)} \sim \hat{d}_{n_t-1,K_C}^{(0)}$. Then, the data at the $(n_t - 2)^{\text{th}}$ transmit antenna can be recovered subsequently. As can be seen from (12), after the $\hat{Q}_{\bar{m}}^H$ multiplication, the signal for

$$\begin{aligned}
x_{\bar{m},n_t-1}^{(1)} &= \hat{h}_{\bar{m},n_t-1}^H \bar{Z}_{\bar{m},n_t-1}^{(1)} \\
&= \left(\sum_{k=0}^{K_C} d_{n_t-1,k} c_{N_F,\bar{m}}^{(k)} \right) \sum_{q=0}^{n_r-1} \left| \hat{h}_{n_t-1,q,\bar{m}} \right|^2 - \underbrace{\left(\sum_{k=0}^{K_C} d_{n_t-1,k} c_{N_F,\bar{m}}^{(k)} \right) \sum_{q=0}^{n_r-1} v_{n_t-1,q,\bar{m}} \hat{h}_{n_t-1,q,\bar{m}}^*}_{\text{from channel estimation error}} \\
&\quad + \underbrace{\sum_{p=0}^{n_t-2} \sum_{q=0}^{n_r-1} \sum_{k=0}^{K_C} \left(d_{p,k} h_{p,q,\bar{m}} - \hat{d}_{p,k}^{(0)} \hat{h}_{p,q,\bar{m}} \right) c_{N_F,\bar{m}}^{(k)} \hat{h}_{n_t-1,q,\bar{m}}^*}_{\text{from residual MTI}} + \underbrace{\sum_{q=0}^{n_r-1} \varepsilon_{q,\bar{m}} \hat{h}_{n_t-1,q,\bar{m}}^*}_{\text{from background noise}} \quad (21)
\end{aligned}$$

the $(n_t - 2)^{\text{th}}$ transmit antenna, $\gamma_{n_t-2,\bar{m}}$, is interfered by data signals from the $(n_t - 1)^{\text{th}}$ transmit antenna. This MTI can be regenerated by using $\hat{d}_{n_t-1,0}^{(0)} \sim \hat{d}_{n_t-1,K_C}^{(0)}$. After MTI cancellation, the resultant signal is further processed and the data at the $(n_t - 2)^{\text{th}}$ transmit antenna are recovered. Similarly, the data at the \bar{p}^{th} transmit antenna can be obtained given the tentative data decisions at the $(\bar{p} + 1)^{\text{th}}$ to $(n_t - 1)^{\text{th}}$ transmit antennas.

2) *Operations in Later Loops:* After the 0^{th} loop, a set of initial data decisions are obtained for all transmit antennas. Since ZF-SIC is employed in the 0^{th} loop, the performance is limited by the signal quality of the first detected antenna, i.e., the $(n_t - 1)^{\text{th}}$ antenna, whose space diversity order is $(n_r - n_t + 1)$. However, as shown in Fig.3, in later loops, MTI cancellation is carried out to detect the signal transmitted on the $(n_t - 1)^{\text{th}}$ antenna. Applying MRC combining, the diversity order of the decision variable is increased to n_r , and the performance should be improved. To be more specific, with the estimated channel information $\hat{H}_{\bar{m}}$ and the tentative data decisions obtained in the 0^{th} loop, i.e., $\hat{d}_{p,k}^{(0)}$ for $p = 0, \dots, n_t - 2$ and $k = 0, \dots, K_C$, the MTI suffered by the signal transmitted on the $(n_t - 1)^{\text{th}}$ antenna can be regenerated as

$$MTI_{\bar{m},n_t-1}^{(1)} = \hat{H}_{\bar{m},n_t-1} \hat{D}_{\bar{m},n_t-1}^{(0)} \quad (19)$$

where $\hat{D}_{\bar{m},n_t-1}^{(0)} = \left(\sum_{k=0}^{K_C} \hat{d}_{0,k}^{(0)} c_{N_F,\bar{m}}^{(k)}, \dots, \sum_{k=0}^{K_C} \hat{d}_{n_t-2,k}^{(0)} c_{N_F,\bar{m}}^{(k)} \right)^T$,

and $\hat{H}_{\bar{m},n_t-1}$ is a $n_r \times (n_t - 1)$ channel matrix with the $(n_t - 1)^{\text{th}}$ column in $\hat{H}_{\bar{m}}$ removed. Define $\bar{D}_{\bar{m},n_t-1}$ as a $(n_t - 1) \times 1$ vector with the $(n_t - 1)^{\text{th}}$ element in $\bar{D}_{\bar{m}}$ removed and $\bar{H}_{\bar{m},n_t-1}$ as a $n_r \times (n_t - 1)$ channel matrix with the $(n_t - 1)^{\text{th}}$ column in $\bar{H}_{\bar{m}}$ removed. After MTI cancellation, the resultant signal is given by

$$\begin{aligned}
\bar{Z}_{\bar{m},n_t-1}^{(1)} &= \bar{Z}_{\bar{m}} - MTI_{\bar{m},n_t-1}^{(1)} \\
&= \left(\sum_{k=0}^{K_C} d_{n_t-1,k} c_{N_F,\bar{m}}^{(k)} \right) \bar{h}_{\bar{m},n_t-1} + \bar{N}_{\bar{m}} \\
&\quad + \underbrace{\bar{H}_{\bar{m},n_t-1} \bar{D}_{\bar{m},n_t-1} - \hat{H}_{\bar{m},n_t-1} \hat{D}_{\bar{m},n_t-1}^{(0)}}_{\text{residual MTI}} \quad (20)
\end{aligned}$$

where the first term contains the desired signal and the third term is the residual MTI. Then the multi-code signal transmitted on the $(n_t - 1)^{\text{th}}$ antenna can be obtained after MRC as shown in (21), where $\hat{h}_{\bar{m},n_t-1} = \bar{h}_{\bar{m},n_t-1} - \bar{v}_{\bar{m},n_t-1}$ is the estimated channel vector, and $\bar{v}_{\bar{m},n_t-1} =$

$(v_{n_t-1,0,\bar{m}}, \dots, v_{n_t-1,n_r-1,\bar{m}})^T$ is the channel estimation error. It can be seen that the useful signal is corrupted by the interferences from $\bar{v}_{\bar{m},n_t-1}$, the residual MTI, and the background noise. Given $x_{\bar{m},n_t-1}^{(1)}$, the hybrid detection can be employed in the frequency domain to recover the data on each code of the $(n_t - 1)^{\text{th}}$ transmit antenna, i.e., $\hat{d}_{n_t-1,0}^{(1)} \sim \hat{d}_{n_t-1,K_C}^{(1)}$.

Generally speaking, with the new data decisions obtained in the 1^{st} loop for the $(\bar{p} + 1)^{\text{th}} \sim (n_t - 1)^{\text{th}}$ transmit antennas, and the recovered data in the 0^{th} loop for the $0^{\text{th}} \sim (\bar{p} - 1)^{\text{th}}$ transmit antennas, the MTI suffered by the signal transmitted from the \bar{p}^{th} antenna can be regenerated and canceled from the received signal. After MRC and hybrid detection, new decisions are produced for the data at the \bar{p}^{th} transmit antenna. The 1^{st} loop of the joint iterative detection is completed as the data symbols at the 0^{th} transmit antenna are recovered. Then the detected data of the 1^{st} loop are fed to the next loop of the joint iterative detection, where the signals are detected in the same way as that of the 1^{st} loop. This iterative operation continues until a specified number of loops is reached.

For an L -loop joint iterative detection with S -stage hybrid detection, the operation at the 0^{th} loop is composed of QR decomposition, S -stage hybrid detection and ZF-SIC. Since the QR decomposition of a $n_r \times n_t$ matrix is needed on M -sub-carriers, the complexity is $O(Mn_t^2 n_r)$ in terms of the number of multiplications. As to the S -stage hybrid detection carried out over every N_F sub-carriers for each transmit antenna, the complexity is given by $O(MSKn_t)$ with M -sub-carriers and n_t transmit antenna. Moreover, the complexity of ZF-SIC is $O(Mn_t n_r)$. Generally, $SK > n_t$. Thus, the complexity of the 0^{th} loop is dominated by that of the QR decomposition and hybrid detection, given by $O(Mn_t^2 n_r + MSKn_t)$. In later loops, the operation is composed of MTI cancellation, MRC and S -stage hybrid detection. With M -sub-carriers, the complexity of MTI cancellation and MRC is of the order of $O(Mn_t^2 n_r)$. Using the complexity of the S -stage hybrid detection derived before, the complexity of later loops is $O(Mn_t^2 n_r + MSKn_t)$. Therefore, the whole complexity of the L -loop joint iterative detection is $O(LMn_t(n_t n_r + SK))$.

IV. OPTIMAL POWER ALLOCATION

In practice, the total transmission power is limited and should be suitably assigned to the data channels and the pilot channel. If too little power is allocated to the pilots, the channel estimation will be inaccurate and the system

$$\begin{aligned}
E \left\{ |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 \right\} &= (n_r - n_t + 1) \cdot E \left\{ |\tilde{h}_{\bar{p}, \bar{q}, \bar{m}}|^2 \right\} \\
&= (n_r - n_t + 1) \cdot \left(P_{ch} \left| \sum_{i=-W}^W \alpha_i \right|^{-2} \sum_{i_0=-W}^W \sum_{i_1=-W}^W \alpha_{i_0} \alpha_{i_1}^* \rho_{i_0, i_1} + \sigma_H^2 \cdot \left| \sum_{i=-W}^W \alpha_i \right|^{-2} \sum_{i=-W}^W |\alpha_i|^2 \right) \\
&= (n_r - n_t + 1) \cdot (P_{ch} C_H + \sigma_H^2 A_H)
\end{aligned} \tag{25}$$

$$\begin{aligned}
P_d \cdot E \left\{ |\hat{r}_{\bar{m}, 0, 0}|^2 \right\} / \sigma^2 &= \frac{P_d (n_r - n_t + 1) (P_{ch} C_H + \sigma_H^2 A_H)}{(K_C + 1) P_d n_t (P_{ch} B_H + \sigma_H^2 A_H) + \frac{\sigma_n^2}{N_T}} \\
&= \frac{(n_r - n_t + 1) (N_T / K) [(2C_H SNR_b (K/N) \cdot N_P N_D + A_H N_P) (\beta / K) + A_H N_D]}{2B_H \cdot SNR_b (K/N) N_P N_D n_t (\beta / K) + ((\beta / K) N_P + A_H n_t) ((\beta / K) N_P + N_D)}
\end{aligned} \tag{27}$$

performance will degrade. On the other hand, if the data channels have low power, they become vulnerable to channel noise and the resultant system performance is also poor. Assume that the transmission power at each antenna is fixed and given by $P_{tot} = MN_D K P_d + MN_P \beta P_d$, then different power allocations are realized by changing the values of β and P_d . When QPSK modulation is employed, the average SNR per bit at each receive antenna is defined as

$$\begin{aligned}
SNR_b &= \frac{P_{ch} N P_{tot}}{2K M N_D \sigma_n^2} = \frac{P_{ch} N M P_d (N_D K + N_P \beta)}{2K M N_D \sigma_n^2} \\
&= \frac{P_{ch} N P_d}{2\sigma_n^2} \left(1 + \frac{N_P \beta}{N_D K} \right)
\end{aligned} \tag{22}$$

The objective of the optimal power allocation is to find the β or β/K that achieves the best system performance for a given SNR_b .

Since the joint iterative detection is based on ZF-SIC, the performance at the firstly detected antenna (or the $(n_t - 1)^{\text{th}}$ transmit antenna) plays an important role to the whole system performance. The proposed optimal power allocation scheme aims to provide the best performance at the $(n_t - 1)^{\text{th}}$ antenna after the 0^{th} stage of hybrid detection (or pure MMSE) in the 0^{th} loop. As described in Sec. III, after multiplying by the $\hat{Q}_{\bar{m}}^H$ matrix, the resultant signal for the $(n_t - 1)^{\text{th}}$ antenna is free of MTI. Then, pure MMSE detection is carried out to collect the desired data from N_F interleaved sub-carriers. Referring to (13), the MSE covariance matrix is given by

$$P_d I_{N_F} - P_d \left[\begin{array}{c} P_d C_F^H H_F^H \\ (\sigma^2 I_{N_F} + P_d H_F C_F C_F^H H_F^H)^{-1} H_F C_F \end{array} \right] \tag{23}$$

whose diagonal elements stand for the MSE on each data code. Assuming a full loaded MIMO-OFCDM system with $K = N$, the diagonal elements are the same and given by

$$\begin{aligned}
MSE &= P_d - \\
&P_d \sum_{\bar{m}=0}^{N_F-1} \left(N_F + \sigma^2 \left(P_d |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 \right)^{-1} \right)^{-1}
\end{aligned} \tag{24}$$

Since the total transmission power at each antenna P_{tot} is fixed and P_d varies with β , the optimal power ratio is obtained by minimizing the normalized MSE, i.e., MSE/P_d , or equivalently, maximizing the term $P_d |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 / \sigma^2$ in (24) for

each \bar{m} , where $|\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2$ follows a chi-square distribution of freedom $2(n_r - n_t + 1)$ [17], with a mean given by (25), where $C_H = \left| \sum_{i=-W}^W \alpha_i \right|^{-2} \sum_{i_0=-W}^W \sum_{i_1=-W}^W \alpha_{i_0} \alpha_{i_1}^* \rho_{i_0, i_1}$. It can be expected that the larger the mean $E \left\{ |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 \right\}$, the higher the probability that $|\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2$ can take a big value. Thus, $P_d \cdot E \left\{ |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 \right\} / \sigma^2$ is taken as the objective function for the optimization instead of $P_d \cdot |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 / \sigma^2$.

First of all, σ^2 can be obtained from (11). Next, P_d is related to β by (see (22))

$$P_d = 2SNR_b \sigma_n^2 / (N + N\beta N_P / (KN_D)) \tag{26}$$

Substituting (25), σ^2 and (26) into the objective function, we obtain (27). (27) has a unique global maximum at β_{opt}/K , which is the solution of the derivative equation $\partial (P_d \cdot E \left\{ |\hat{r}_{\bar{m}, n_t-1, n_t-1}|^2 \right\} / \sigma^2) / \partial (\beta/K) = 0$, given by

$$\begin{aligned}
\beta_{opt}/K &= - \left(\frac{B}{A} \right) \\
&+ \sqrt{\left[A_H \left(1 - \frac{B_H}{C_H} \right) \frac{n_t}{N_P} - \left(\frac{B}{A} \right) \right] \left[\frac{N_D}{N_P} - \left(\frac{B}{A} \right) \right]}
\end{aligned} \tag{28}$$

where $A_{SNR} = 2SNR_b (K/N) (N_D/N_P)$ and $B/A = A_H N_D / (A_H N_P + C_H N_P^2 \cdot A_{SNR})$. As SNR_b approaches infinity, B/A comes near to zero, and (28) can be simplified to

$$\lim_{SNR_b \rightarrow +\infty} \beta_{opt}/K = \sqrt{A_H (1 - B_H/C_H) n_t N_D / N_P^2} \tag{29}$$

It can be seen from (28) that the derived β_{opt}/K only depends on SNR and the properties of the channel estimation represented by A_H , B_H and C_H , the spatial efficiency n_t/N_P and the time efficiency N_D/N_P . This is because the pilot channels at different antennas are orthogonal to each other. The channel estimation is free of MTI or MDI and only corrupted by background noise. Moreover, β_{opt}/K increases with SNR. When SNR is small, the system performance is dominated by the channel noise. More power should be allocated to data channels to combat the background noise. As SNR increases, the effect of channel noise reduces and

TABLE I
SYSTEM PARAMETERS

No. of transmit antennas n_t	4
No. of receive antennas n_r	4
Bandwidth	100MHz
No. of sub-carriers M	1024 (97.7kHz sub-carrier spacing)
OFCDM symbol duration $T_s = T_e + T_g$	$12.5\mu s = 10.24 + 2.26$
Packet length ($T_s \cdot (N_P + N_D)$) (No. of OFCDM symbols per packet)	0.65ms ($12.5\mu s \cdot (4 + 48)$) (52 OFCDM symbols: 4 for pilot and 48 for data)
Pilot Channel	Time-multiplexed with data channels, one dimensional spreading in the time domain
Spreading Code for Data Channel	2-D OVFS (spreading factor $N = N_T \times N_F = 4 \times N_F$)
Spreading Code for Pilot Channel	OVFS (spreading factor $N_P = 4$)
System Load	$K/N = 1.0$
Modulation	QPSK for both data and pilot channels
Channel Model	Parallel slow fading multi-channel model with a coherence bandwidth of $\Delta f_c = 1\text{MHz}$
Channel Estimation	Pilot-aided, 2-D averaging channel estimation ($W = 1$, $\alpha_0 = 1$, $\alpha_{-1} = \alpha_1 = 0.8$)
Signal Detection	4-Loop joint iterative detection with 2-stage hybrid detection

the channel estimation error plays a more important role in determining the system performance. Thus, more power is allocated to the pilot channel to obtain accurate channel information. When SNR is sufficiently large, as shown in (29), β_{opt}/K is decided by the properties of the channel estimation and system parameters n_t , N_P , and N_D .

It can also be seen that the derived β_{opt}/K is not related to the number of receiving antennas n_r or the frequency domain spreading factor N_F since n_r and N_F have no influence on the channel estimation quality or the amount of channel noise experienced by data channels. However, the system performance improves as n_r and N_F increase due to the larger space and frequency domain diversity, respectively. Moreover, the derived β_{opt}/K aims to provide the best performance at the $(n_t - 1)^{\text{th}}$ antenna after the 0^{th} stage of hybrid detection (or pure MMSE) in the 0^{th} loop. Yet using the joint iterative detection, the whole system performance is much better than that in the 0^{th} loop. So the system is more robust to the background noise than that with MMSE detection and more power than that indicated by (28) should be allocated to the pilot channel, especially at low SNR. Hence, (28) provides a lower bound for the β_{opt}/K achieving the best system performance. On the other hand, by ignoring the effect of background noise, (29) is an upper bound for the β_{opt}/K . In fact, as verified in the next section, (29) provides an accurate estimation for the β_{opt}/K with various system conditions.

V. PERFORMANCE EVALUATION

The configuration of the MIMO-OFCDM system is shown in Table I. Using these parameters, the total number of infor-

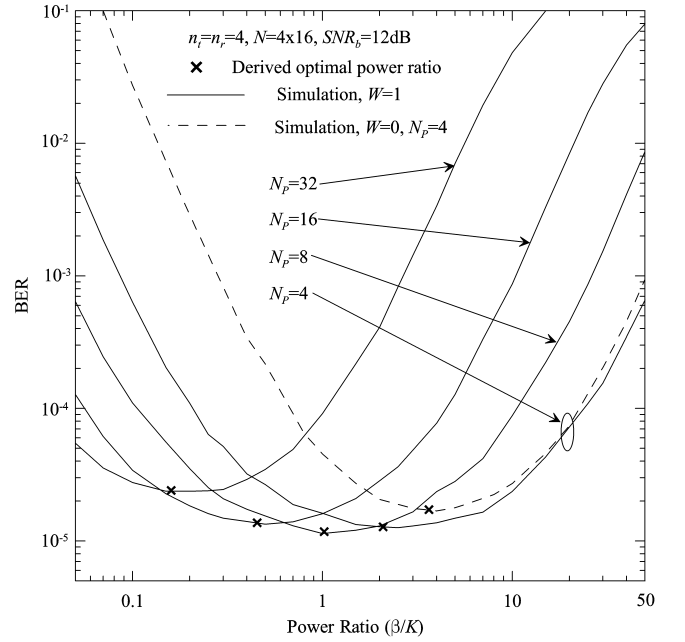
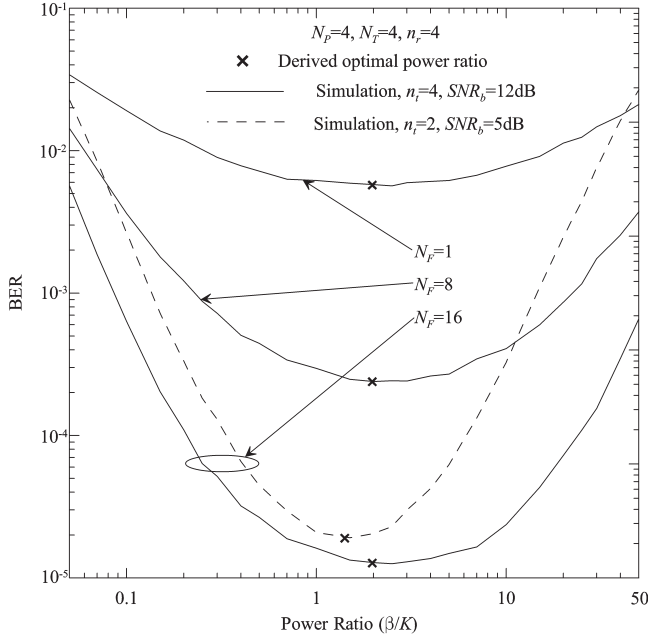


Fig. 4. Verification of the optimal power ratio with different N_P

mation bits transmitted in one packet is $2M \cdot N_D \cdot K/N \approx 10^5$. Given Δf_c , the correlated Rayleigh fading channels can be generated according to [23]. When evaluating the quality of channel estimation, the parameters A_H , B_H and C_H are calculated as $A_H = 0.34$, $B_H = 7.42e - 5$ and $C_H = 0.99$. Thus, the power ratio between the pilot channel and all data channels can be obtained from (29) as $\beta/K = 2.02$. All simulation results are obtained after at least 1000 bit errors are collected.

A. Verify Derived Optimal Power Ratio

First of all, the derived optimal power ratio (see (29)) is verified by simulation results. Given $n_t = n_r = 4$ and $N_P + N_D = 52$, the average BER of the recovered symbols at four transmit antennas is shown in Fig.4 as a function of β/K with different number of pilot symbols N_P . In general, it can be seen that when β/K is small, the pilot channel has small power and the channel estimation quality is poor. Thus, the system performance is poor and the average BER is high. When β/K increases, the BER reduces as the quality of channel estimation improves. The BER reaches a minimum value for a particular value of β/K . Further increasing β/K beyond that value increases BER due to too little power assigned to data channels. With $N_P = 32, 16, 8$, and 4 , the minimum average BER is achieved when β/K is around 0.15, 0.45, 1.0 and 2.0, respectively. Correspondingly, using (29), the analytical optimal power ratios can be calculated as 0.16, 0.44, 0.97 and 2.02, respectively, which match well with the simulations. Moreover, given $N_P = 4$, simulation results are also shown for the system using the channel estimation algorithm with $W = 0$ and only time domain averaging is used. In this case, $A_H = 1.0$, $B_H = 0$, $C_H = 1.0$, and the derived $\beta_{opt}/K = 3.46$. Again, the analytical result coincides with the simulations. It can be seen that the optimal power ratio decreases from 3.5 to 2.0 when W increases from zero

Fig. 5. Verification of the optimal power ratio with different N_F and n_t

to one. Because the channel estimation is more accurate with 2-D averaging ($W = 1$) than that with only time averaging ($W = 0$), the system performance is improved and less power is needed on the pilot channel. Furthermore, using $W = 1$, the system is more robust to the changes in β/K since near optimum performance can be achieved when β/K varies between 1.0 and 7.0. Therefore, the 2-D averaging channel estimation is preferred to the time averaging one not only due to better system performance but also due to the increased robustness to the changes in β/K . Moreover, although the analytical optimal power ratio is derived by minimizing the MSE on the $(n_t - 1)^{\text{th}}$ antenna after the 0^{th} stage of hybrid detection in the 0^{th} loop, it turns out that after 4-loop joint iterative detection, low BERs are achieved for all antennas as well.

Next, given $N_P = 4$ and $n_r = 4$, the average BER is shown in Fig.5 as a function of β/K with different N_F and n_t . First of all, it can be seen that the system performance is greatly enhanced when N_F increases from 1 to 8 and 16 due to the increased frequency diversity gain. But the near optimum performance is always achieved when β/K takes values around 2.0 and matches with the analysis. On the other hand, when $SNR_b = 5\text{dB}$, $n_t = 2$ and $n_r = 4$, the system obtains a receive diversity gain and achieves the best performance when β/K is around 1.5. Using (28) and (29), the β_{opt}/K can be calculated as 1.35 and 1.43, respectively, both matching with the simulation. Considering the increased robustness to channel noise due to the receive diversity gain, more power than that indicated by (28) should be allocated to the pilot channel. Thus, (29) is a more appropriate estimation for the β_{opt}/K . In summary, it has been verified that the β_{opt}/K obtained from (29) works well for the system with various frequency diversity or receive diversity gain. Moreover, it is also illustrated in Fig.6 that the derived β_{opt}/K matches with the simulations with different system parameters such as

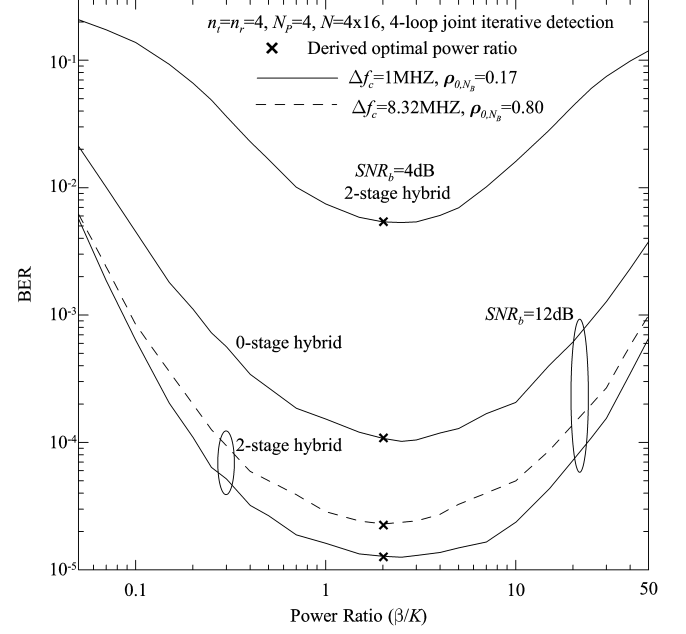


Fig. 6. Verification of the optimal power ratio with other system parameters

SNR , sub-carrier correlations and number of stages in the joint iterative detection.

In summary, the simulation is consistent with (29) that the optimal power ratio is mainly determined by the quality of channel estimation, while SNR , diversity gains, sub-carrier correlations and detection algorithms have little effect on the value of β_{opt}/K . The derived optimal power ratio (29) is shown to be accurate with various system configurations and channel conditions. In the rest of the paper, n_t and N_P are fixed to 4 and β/K is set to 2.0 which is a near optimum power ratio.

B. Performance of Joint Iterative Detection

Given $SNR_b = 12\text{dB}$ and $N = 4 \times 16$, the effect of the number of loops in the joint iterative detection is investigated in Fig.7. As a comparison, the BER of the OFCDM system with single antenna (SISO-OFCDM) is also plotted, where only the hybrid detection can be employed to exploit the frequency domain diversity and there is no space diversity. Similarly, the BER of the MIMO-OFCDM system without frequency domain spreading (i.e., $N_F = 1$) is also shown, where only the iterative detection in the space domain can be employed and there is no frequency diversity. It can be seen that for SISO-OFCDM or the MIMO-OFCDM with $N_F = 1$, the iterative detection can improve the system performance, but the improvement is limited. However, when the joint iterative detection is employed in the MIMO-OFCDM with $N = 4 \times 16$, the performance is enhanced significantly when the number of loops increases. This is due to the fact that using the proposed detection, the MIMO-OFCDM can jointly exploit the diversity gains in both space and frequency domains. A second observation is that from the 0^{th} to 1^{st} loop in the joint iterative detection, the BER reduction is significant. The improvement in BER between the 1^{st} loop and 2^{nd} loop is also considerable. However, the improvement is insignificant

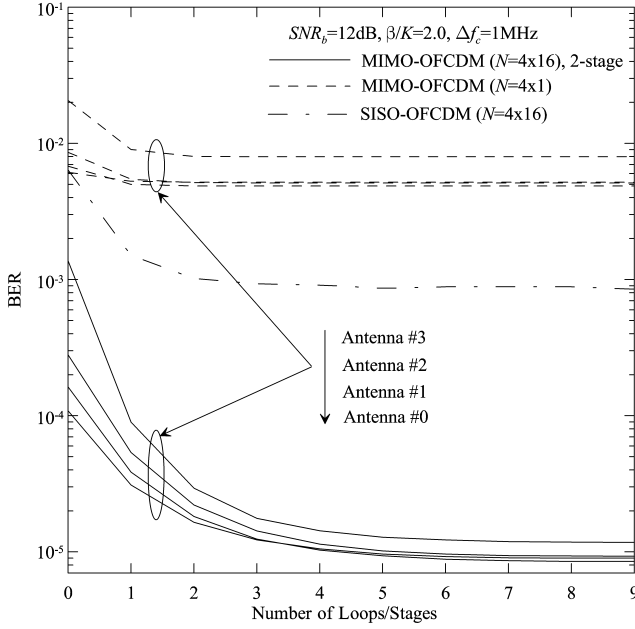


Fig. 7. Effect of number of loops in joint iterative detection

beyond the 4th loop. Finally, it can also be seen that the BER differences among the four antennas reduce as the number of loops increases. The reason is that in the 0th loop, ZF-SIC is employed in the space domain which provides no space diversity gain for the $(n_t - 1)^{th}$ antenna while later detected antennas can obtain greater space diversity. Therefore, the BER at the 0th antenna is much lower than that at the $(n_t - 1)^{th}$ antenna. However, in later loops, MTI cancellation and MRC are carried out in the space domain. In this case, signals at all antennas achieve similar space diversity gain and the BERs on all antennas are therefore close to each other. As a conclusion, for the MIMO-OFCDM, a 4-loop joint iterative detection with 2-stage hybrid detection is sufficient to provide good performance.

The performance of MIMO-OFCDM is illustrated in Fig.8 as a function of SNR. The average BER on four transmitting antennas is shown for each loop. It is shown that the BER improves as the number of loops increases, especially at larger SNR. The BER gap between the 0th and 1st loop is considerable, while for later loops, the performance improvement reduces. This is consistent with that in Fig.7.

Fig.9 illustrates the effect of sub-carrier correlation ρ_{0,N_B} on the BER performance, where $N_B = M/N_F$ and ρ_{0,N_B} is the correlation between adjacent sub-carriers carrying the same data symbol. Given a system bandwidth, the number of sub-carriers M and spreading scheme $N = N_T \times N_F$, the correlation coefficient ρ_{0,N_B} varies with channel coherence bandwidth Δf_c . It can be seen that when ρ_{0,N_B} is very low, e.g., less than 0.1, the BER is high. The reason is that the system employs a 2-D averaging channel estimation algorithm with $W = 1$ and $\alpha_{-1} = \alpha_1 = 0.8$, which become inappropriate for an extremely frequency selective channel with very low ρ_{0,N_B} . When ρ_{0,N_B} increases, the BER reduces as better channel estimation is obtained. As a comparison, the BER with $W = 0$ is also shown for the 4th loop. It can be seen that without frequency domain averaging, the channel estimation

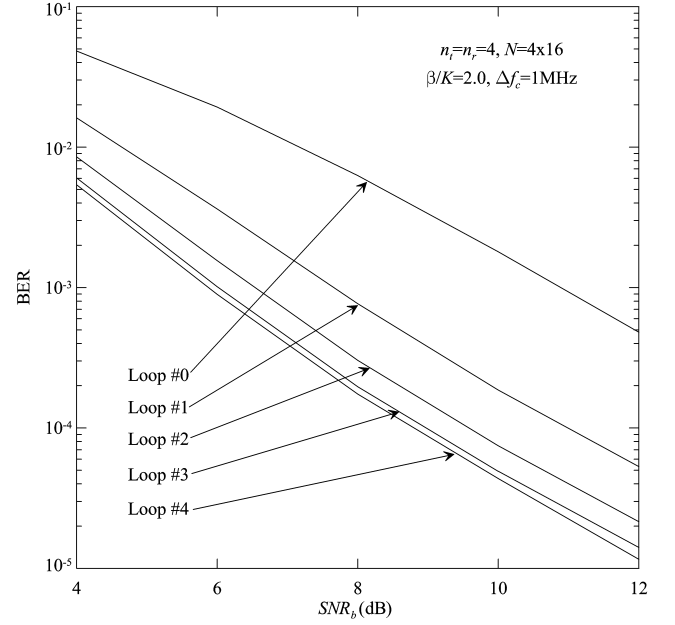


Fig. 8. Performance of joint iterative detection as a function of SNR

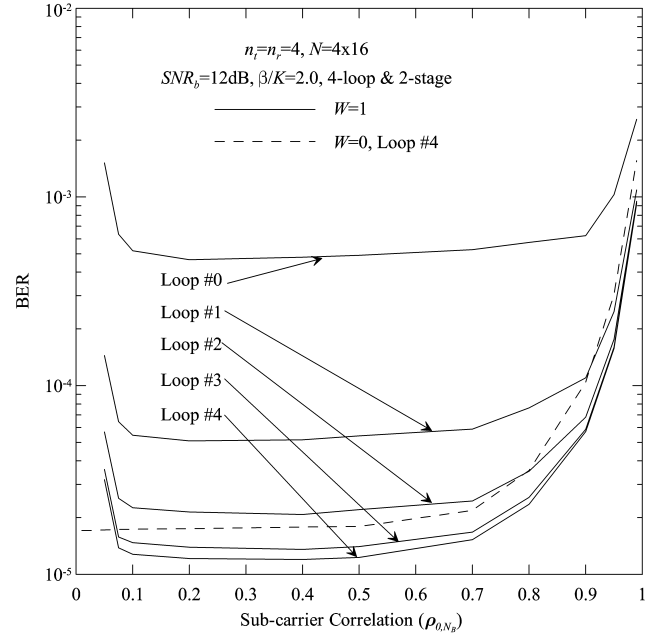


Fig. 9. System performance as a function of sub-carrier correlations

quality is not affected by the sub-carrier correlation and the system performance keeps stable at low ρ_{0,N_B} . Although the 2-D average channel estimation with fixed parameters is not as good as the one with only time domain averaging at low ρ_{0,N_B} , it works well for most correlation conditions. On the other hand, when ρ_{0,N_B} increases further, the frequency diversity gain obtained from the frequency domain spreading decreases, resulting in a degraded BER. However, the BER degradation is negligible for $\rho_{0,N_B} \leq 0.7$.

Finally, the performance of MIMO-OFCDM is shown in Fig.10 as a function of the frequency domain spreading factor N_F . When N_F increases, the performance of MIMO-

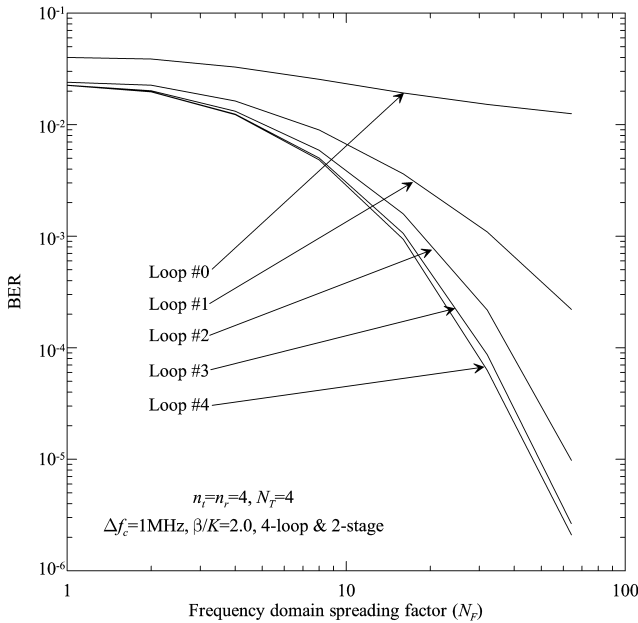


Fig. 10. Effect of frequency domain spreading factor

OFCDM improves. Since the system load K/N is fixed, the performance enhancement resulted from a larger N_F is not due to the spreading gain, but due to the increased spatial and frequency diversity gain. In the 0th loop, the BER is flat and the improvement is minimal because the system performance is limited by the $(n_t - 1)^{\text{th}}$ antenna which has no space diversity gain. However, in later loops, there is significant performance improvement when N_F increases since each antenna can benefit from both space and frequency diversity. When N_F becomes larger, the tentative data decisions in previous loops are more precise, and the MTI is regenerated with higher accuracy. After MTI cancellation and MRC combining, more space diversity gain can be obtained. Meanwhile, when N_F increases, the frequency spacing between the sub-carriers carrying the same data symbol is decreased and the corresponding correlation ρ_{0,N_B} is increased. Given a fixed coherence bandwidth of $\Delta f_c = 1\text{MHz}$, when N_F increases from 1 to 64, ρ_{0,N_B} increases from 0.0 to 0.54. However, as illustrated in Fig.9, only when the sub-carrier correlation is larger than 0.7, the degradation in BER performance will be obvious. Therefore, as a whole, the system performance improves when N_F increases from 1 to 64. However, it can be expected that the improvement in system performance would become marginal if N_F increases further.

VI. CONCLUSION

A joint iterative detection algorithm has been proposed for the MIMO-OFCDM system in this paper. A close form expression has been derived for the optimal power ratio (β_{opt}/K) to achieve the best system performance. It has been shown that β_{opt}/K is mainly decided by the channel estimation algorithm, the number of transmit antennas as well as the number of pilot and data symbols in a packet, while it is not sensitive to the changes in SNR and diversity gains. The system performance has been investigated extensively by simulations. Under various channel conditions, the proposed

joint iterative detection algorithm significantly enhances the system performance by jointly exploiting the space and frequency diversity gains. Considerable improvement in BER could be obtained when the number of loops increases but with diminishing returns. The system performance also improves significantly as the frequency domain spreading factor, N_F , increases.

ACKNOWLEDGMENT

The authors would like to thank the Editor, M. Morelli, as well as the anonymous reviewers for their careful reading and valuable comments that greatly improve the paper.

REFERENCES

- [1] G. J. Foschini, Jr., "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41-59, Autumn 1996.
- [2] R. D. Murch and K. B. Letaief, "Antenna systems for broadband wireless access," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 76-83, Apr. 2002.
- [3] A. Sumasu, T. Nihei, K. Kitagawa, M. Uesugi, and O. Kato, "An OFDM-CDMA system using combination of time and frequency domain spreading," *Technical Report of IEICE*, RCS2000-3, Apr. 2000.
- [4] A. Persson, T. Ottosson, and E. Strom, "Time-frequency localized CDMA for downlink multi-carrier systems," in *Proc. IEEE ISSSTA 2002*, pp. 118-122, Sept. 2002.
- [5] H. Atarashi, S. Abeta, and M. Sawahashi, "Broadband packet wireless access appropriate for high-speed and high-capacity throughput," in *Proc. IEEE VTC2001-Spring*, pp. 566-570, May 2001.
- [6] *Multiplexing and Channel Coding (FDD)*, 3rd Generation Partnership Project, Tech. Spec. Group Radio Access Network, ETSI TS 25212 V3.5.0 (Rel. 1999), 2000.
- [7] N. Maeda, T. Kataoka, H. Kawai, K. Higuchi, J. Kawamoto, and M. Sawahashi, "Experiments on real-time 1-Gbps packet transmission using antenna-independent AMC in MIMO-OFDM broadband packet radio access," in *Proc. IEEE VTC-2005-Fall*, vol. 3, pp. 1628-1632, Sept. 2005.
- [8] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-SS-CDMA for next-generation mobile communications systems," *IEEE Commun. Mag.*, vol. 36, no. 9, pp. 56-69, Sept. 1998.
- [9] B. Xia, J. Wang, and M. Sawahashi, "Performance comparison of optimum and MMSE receivers with imperfect channel estimation for VSF-OFCDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 3051-3062, Nov. 2005.
- [10] Y. Q. Zhou, J. Wang, and M. Sawahashi, "Downlink transmission of broadband OFCDM systems—part I: hybrid detection," *IEEE Trans. Commun.*, vol. 53, pp. 718-729, Apr. 2005.
- [11] —, "Downlink transmission of broadband OFCDM systems—part III: turbo-coded," *IEEE J. Select. Areas Commun.*, vol. 24, no. 1, pp. 132-140, Jan. 2006.
- [12] K. Higuchi, H. Kawai, N. Maeda, and M. Sawahashi, "Adaptive selection of surviving symbol replica candidates based on maximum reliability in QRM-MLD for OFCDM MIMO multiplexing," in *Proc. IEEE Globecom2004*, pp. 2480-2486, Nov. 2004.
- [13] H. Kawai, K. Higuchi, N. Maeda, M. Sawahashi, T. Ito, Y. Kakura, A. Ushirokawa, and H. Seki, "Likelihood function for QRM-MLD suitable for soft-decision turbo decoding and its performance for OFCDM MIMO multiplexing in multipath fading channel," *IEICE Trans. Commun.*, vol. E88-B, no. 1, pp. 47-57, Jan. 2005.
- [14] W. Yin, C. C. Tsimenidis, and B. S. Sharif, "Transmission power optimization of convolutional coded VBLAST system," in *Proc. IEEE VTC2005-Spring*, pp. 945-948, June 2005.
- [15] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich scattering wireless channel," in *Proc. IEEE ISSSE-98*, pp. 295-300, Sept. 1998.
- [16] D. Wbenn, R. Bohnke, J. Rinas, V. Khn and K. D. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *Electron. Lett.*, vol. 37, pp. 1348-1350, Oct. 2001.
- [17] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: an analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1326-1337, July 2004.
- [18] C. Shen, H. Zhuang, L. Dai, and S. Zhou, "Detection algorithm improving V-BLAST performance over error propagation," *IEE Electron. Lett.*, vol. 39, pp. 1007-1008, June 2003.

- [19] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, pp. 3065-3073, Dec. 2001.
- [20] W. C. Y. Lee, *Mobile Communications Engineering: Theory and Applications*, 2nd edition. New York: McGraw-Hill, 1997.
- [21] H. Kawai, K. Higuchi, N. Maeda, and M. Sawahashi, "Performance of QRM-MLD employing two-dimensional multi-slot and sub-carrier averaging channel estimation filter using orthogonal pilot channel for OFCDM MIMO multiplexing in multipath fading channel," in *Proc. Wireless 2004*, pp. 208-214, July 2004.
- [22] Y. Q. Zhou, J. Wang, T. S. Ng, K. Higuchi, and M. Sawahashi, "OFCDM: a promising broadband wireless access technique," *IEEE Commun. Mag.*, vol. 46, pp. 39-49, Mar. 2008.
- [23] 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.



Yiqing Zhou (S'03, M'05) received the B.S. degree in communications and information engineering and the M.S. degree in signal and information processing from the Southeast University, Nanjing, China, in 1997 and 2000, respectively. In February 2004, she received the Ph.D. degree in electrical and electronic engineering from the University of Hong Kong, Hong Kong. Since June 2004, she has been with the Department of Electrical and Electronic Engineering at the University of Hong Kong as a Postdoctoral Fellow. Her research interests include coding theory,

spread spectrum, OFDM systems, interference cancellation, hybrid ARQ, MIMO and other transmission techniques for wireless high speed data communications.



Tung-Sang Ng received the B.Sc.(Eng.) degree from The University of Hong Kong in 1972, the M.Eng.Sc. and Ph.D. degrees from the University of Newcastle, Australia, in 1974 and 1977, respectively, all in electrical engineering.

He worked for BHP Steel International and The University of Wollongong, Australia after graduation for 14 years before returned to The University of Hong Kong in 1991, taking up the position of Professor and Chair of Electronic Engineering. He was Head of Department of Electrical and Electronic

Engineering from 2000 to 2003 and Dean of Engineering from 2003 to 2007. His current research interests include wireless communication systems, spread spectrum techniques, CDMA and digital signal processing. He has published over 250 international journal and conference papers.

He was the General Chair of ISCAS'97 and the VP-Region 10 of IEEE CAS Society in 1999 and 2000. He was an Executive Committee Member and a Board Member of the IEE Informatics Divisional Board (1999-2001) and was an ordinary member of IEE Council (1999-2001). He was awarded the Honorary Doctor of Engineering Degree by the University of Newcastle, Australia in 1997, the Senior Croucher Foundation Fellowship in 1999, the IEEE Third Millenium medal in 2000 and the Outstanding Researcher Award by The University of Hong Kong in 2003. He is a Fellow of IEEE, IEE and HKIE.