PAPER Special Section on Smart Multimedia & Communication Systems

An Improved Channel Shortening Method with Application to MC-CDMA Systems

Mizuki KOTAKE[†a)], ***Student Member and*** **Teruyuki MIYAJIMA**[††], ***Member***

SUMMARY In block transmissions, inter-block interference (IBI) due to delayed waves exceeding a cyclic prefix severely limits the performance. To suppress IBI in downlink MC-CDMA systems, this paper proposes a novel channel shortening method using a time-domain equalizer. The proposed method minimizes a cost function related to equalizer output autocorrelations without the transmission of training symbols. We prove that the method can shorten a channel and suppress IBI completely. Simulation results show that the proposed method can significantly suppress IBI using relatively less number of received blocks than a conventional method when the number of users is moderate.

key words: *block transmissions, multi-carrier code-division multiple-access, blind equalization, inter-block interference*

1. Introduction

Recent progress in digital broadcasting, wireless LAN and cellular phone systems demands techniques to process a larger amount of data at higher speed. Block transmissions have been valuable candidates in these applications. Orthogonal frequency division multiplexing (OFDM) is a well-known single-user modulation technique that has advantages in high data rate transmission in frequency selective channels [1]. Orthogonal frequency division multiple access (OFDMA) as an extension of OFDM allows multiple users to access a wideband channel dynamically, and has been widely adopted in several standards [2]. Multicarrier code-division multiple-access (MC-CDMA) [3] is a combination of OFDM with CDMA for supporting multiple access communications and has been attracted much attention as a broadband wireless access technique [4]–[7].

In block transmissions, if the length of a channel impulse response is longer than the length of a cyclic prefix (CP), which is inserted between blocks, residual inter-block interference (IBI) severely limits the system performance. Several IBI suppression techniques have been addressed in the literatures, such as nonlinear interference cancelers [8], adaptive antenna arrays [9] and channel shortening using time-domain equalizers (TEQ) [10], [11]. This paper focuses on channel shortening methods that shorten the impulse response of the effective channel composed of a physical channel and a TEQ to within the CP length to suppress IBI.

Most of existing channel shortening methods require the transmission of training signals that waste bandwidth [10]. Meanwhile, blind methods without training signals are attractive because of their high spectral efficiency [12]–[17]. A pioneering work of blind channel shortening is Multicarrier Equalization by Restoration of RedundancY (MERRY) algorithm [12] that has been successfully applied to single-user OFDM systems. The MERRY algorithm, as well as most of other blind shortening methods, assumes that the transmitted time-domain samples are uncorrelated. In practice, however, the transmitted samples can be highly correlated [11]. For example, in OFDMA systems, the existence of null subcarriers causes correlation among transmitted samples; in MC-CDMA systems, a number of subcarriers convey a common data symbol, and the resulting transmitted samples are correlated with each other.

A recent work in [17] has investigated channel shortening methods for OFDMA systems, and demonstrated the potential suitability of the carrier nulling algorithm [14]. On the other hand, there has been only limited work in [16] addressing channel shortening for MC-CDMA systems. In [16], it has been proved that the MERRY algorithm cannot shorten a channel, while a shortening method based on second-order statistics of the received signals can shorten a channel in a single-user MC-CDMA system. Simulation results, however, revealed that the latter method requires a large number of received blocks to acquire precisely the statistics. Therefore, we need an improved channel shortening method that requires a less number of received blocks in multiuser environments.

This paper proposes an improved blind channel shortening method applicable to MC-CDMA systems. In the proposed method, TEQ is determined by minimizing a cost function related to equalizer output autocorrelations. We prove that the proposed method can shorten the impulse response of an effective channel to within the CP length under certain conditions. In simulations, we compare the performance of the proposed method with that of conventional methods, and discuss pros and cons of these methods.

2. System Model

A baseband discrete-time downlink MC-CDMA system consisting of K users and N subcarriers is shown in Fig. 1. The kth user's data sequence $\{d_n^{(k)}\}$ is generated at a rate $1/T_s$. The following assumption on $\{d_n^{(k)}\}$ is made:

Manuscript received January 24, 2012.
Manuscript revised June 1, 2012.
[†]The author is with the Graduate School of Sci. and Eng., Ibaraki University, Hitachi-shi, 316-8511 Japan.
[††]The author is with the Faculty of Eng., Ibaraki University, Hitachi-shi, 316-8511 Japan.
a) E-mail: 11nm615t@hcs.ibaraki.ac.jp
DOI: 10.1587/transfun.E95.A.1955

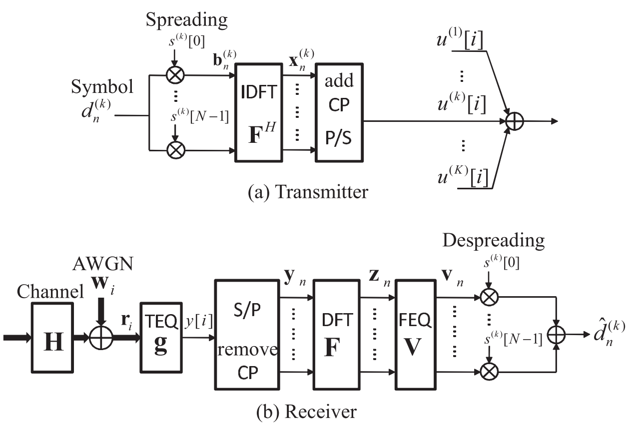

Fig. 1 Downlink MC-CDMA system with a TEQ.

AS1) $\{d_n^{(k)}\}$ is i.i.d with a variance σ_d^2 and uncorrelated with $\{d_n^{(l)}\}, l \neq k$.

The kth MC-CDMA user spreads its nth data symbol $d_n^{(k)}$ over different subcarriers using a preassigned spreading code in the frequency domain [3]. Let us denote the spreading code assigned to the kth user as $\mathbf{s}^{(k)} = [s^{(k)}[0] \cdots s^{(k)}[N-1]]^T$ where $s^{(k)}[i] \in \{\pm 1\}$ and the superscript T represents the transpose of a matrix or vector. We employ orthogonal spreading codes $\mathbf{s}^{(k)}$ that satisfy

$$\mathbf{s}^{(k)T}\mathbf{s}^{(l)} = \begin{cases} N & l = k \\ 0 & l \neq k \end{cases}. \tag{1}$$

Then we can write the nth block of chip sequence $\mathbf{b}_n^{(k)}$ corresponding to the data symbol $d_n^{(k)}$ as

$$\mathbf{b}_n^{(k)} = d_n^{(k)}\mathbf{s}^{(k)}. \tag{2}$$

The $N \times N$ DFT matrix expressed by $\mathbf{F}$ whose (m,n)th element is $\exp(j2\pi mn/N)/\sqrt{N}$ where $m, n = 0, \cdots, N-1$. Applying N-point IDFT $\mathbf{F}^H$ to $\mathbf{b}_n^{(k)}$ gives a time-domain block vector

$$\mathbf{x}_n^{(k)} = \mathbf{F}^H\mathbf{b}_n^{(k)} \tag{3}$$

where the superscript H represents the Hermitian transpose of a matrix or vector. The CP of length P is added to the beginning of $\mathbf{x}_n^{(k)}$ to obtain the transmitted samples $u^{(k)}[i], i = nQ, \cdots, (n+1)Q-1$, where $Q = N + P$ is the length of a transmitted block.

Clearly, an element of $\mathbf{x}_n^{(k)}$ is correlated with the other elements of $\mathbf{x}_n^{(k)}$. As mentioned later, this nonzero correlation makes shortening problem to be difficult. It should be noted that each element of $\mathbf{x}_n^{(k)}$ is uncorrelated with elements of $\mathbf{x}_m^{(k)}, m \neq n$, because of the i.i.d. assumption on $\{d_n^{(k)}\}$, that is, the transmitted samples of a block are uncorrelated with those of different blocks. This property plays an important role to develop a novel shortening method.

We consider downlink MC-CDMA systems with D receive antennas. The received signal of the dth antenna is given by

$$r_d[nQ+i] = \sum_{k=1}^{K}\sum_{m=0}^{M} h_d[m]u^{(k)}[nQ+i-m] + w_d[nQ+i] \tag{4}$$

for $i = 0, \cdots, Q-1$ and $d = 0, \cdots, D-1$, where $\{h_d[i]\}$ is the impulse response of the dth downlink channel, M is the order of the channel and $\{w_d[i]\}$ represents channel noise that is a stationary spatially white Gaussian noise with variance σ_w^2 and uncorrelated with the data sequences.

Let us consider a time-domain equalizer (TEQ) that is an FIR filter of length L. The TEQ output is given by

$$y[i] = \sum_{d=0}^{D-1}\sum_{l=0}^{L-1} g_d^*[l]r_d[i-l] = \mathbf{g}^H\mathbf{r}_i \tag{5}$$

where $\mathbf{g}$ is the TEQ vector of length LD, and the ith received sample vector of length LD can be written as [15]

$$\mathbf{r}_i = \mathbf{H}\sum_{k=1}^{K}\mathbf{u}_i^{(k)} + \mathbf{w}_i = \mathbf{H}\mathbf{u}_i + \mathbf{w}_i \tag{6}$$

where $\mathbf{w}_i$ is a noise vector of length LD, the ith sampled channel input vector of length $L+M$ can be expressed as $\mathbf{u}_i^{(k)} = [u^{(k)}[i] \cdots u^{(k)}[i-(L+M)+1]]^T$, an $LD \times (L+M)$ channel matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}[0] & \cdots & \mathbf{h}[M] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}[0] & \cdots & \mathbf{h}[M] \end{bmatrix} \tag{7}$$

where $\mathbf{h}[i] = [h_0[i] \cdots h_{D-1}[i]]^T$, and $\mathbf{u}_i = \sum_{k=1}^{K}\mathbf{u}_i^{(k)}$. Henceforth, the following assumptions are made:

AS2) $h_d[i]$ can be modeled as a complex Gaussian variable, and $\mathbf{H}$ has full column rank,

and $L + M \leq N$. We further assume that the channel is quasi static so that the path gains are constant during a long period of time. The TEQ output can be rewritten as

$$y[i] = \mathbf{c}^H\sum_{k=1}^{K}\mathbf{u}_i^{(k)} + \mathbf{g}^H\mathbf{w}_i, \tag{8}$$

where $\mathbf{c}$ is an $(L+M) \times 1$ vector

$$\mathbf{c} = [c[0] \cdots c[L+M-1]]^T = \mathbf{H}^H\mathbf{g} \tag{9}$$

that denotes the impulse response of the effective channel that is a convolution of a physical channel $\{h_d[i]\}$ and TEQ $\{g_d[i]\}$. The purpose of channel shortening is to find $\mathbf{g}$ such that $\{c[i]\}$ is shorten to within $P+1$, i.e.,

$$c[i] = 0, \; i = P+1, \cdots, L+M-1. \tag{10}$$

We can write TEQ outputs after removing CP in a vector form as $\mathbf{y}_n = [y[nQ+P] \cdots y[(n+1)Q-1]]^T$. Applying the N-point DFT $\mathbf{F}$ to $\mathbf{y}_n$, we get the DFT output vector as

$$\mathbf{z}_n = \mathbf{F}\mathbf{y}_n. \tag{11}$$

Frequency-domain equalizer (FEQ) can be represented by a diagonal matrix $\mathbf{V}$. The output of FEQ is given by

$$\mathbf{v}_n = \mathbf{V}\mathbf{z}_n. \tag{12}$$

Finally, we get an estimate of $d_n^{(k)}$ by de-spreading the FEQ output

$$\hat{d}_n^{(k)} = \mathbf{s}^{(k)T}\mathbf{v}_n. \tag{13}$$

If the effective channel $\{c[i]\}$ is shortened to within $P+1$, i.e., IBI is suppressed completely, and noise is sufficiently small, the DFT output becomes

$$\mathbf{z}_n = \mathbf{Q}\sum_{k=1}^{K} d_n^{(k)}\mathbf{s}^{(k)} \tag{14}$$

where $\mathbf{Q}$ is a diagonal matrix whose ith diagonal element is $q[i]$ where $q[i]$ is the frequency gain at ith tone of the effective channel $\mathbf{c}$. We can see from (14) that inter-carrier interference (ICI) does not exist in the DFT outputs. Then, if we use zero-forcing (ZF) FEQ $\mathbf{V} = \mathbf{Q}^{-1}$, the kth user's estimated symbol becomes

$$\hat{d}_n^{(k)} = \mathbf{s}^{(k)T}\mathbf{V}\mathbf{Q}\sum_{l=1}^{K} d_n^{(l)}\mathbf{s}^{(l)T} = \mathbf{s}^{(k)T}\sum_{l=1}^{K} d_n^{(l)}\mathbf{s}^{(l)} = N d_n^{(k)}. \tag{15}$$

Then multiple-access interference (MAI) is also cancelled.

3. Conventional Shortening Methods

3.1 MERRY Algorithm

Let us discuss the applicability of MERRY to multiuser MC-CDMA systems. The MERRY algorithm tries to retrieve the inherent CP induced cyclostationarity that is disturbed by multipath propagation, and minimizes the following cost function subject to the norm constraint of $\|\mathbf{g}\| = 1$:

$$J_0(\mathbf{g}) = E[|y[nQ+P-1] - y[nQ+Q-1]|^2]. \tag{16}$$

The cost function is also written by a matrix form as $J_0(\mathbf{g}) = \mathbf{g}^H\bar{\mathbf{R}}\mathbf{g}$ where an $LD \times LD$ matrix $\bar{\mathbf{R}}$ is given by $\bar{\mathbf{R}} = E[(\mathbf{r}_{nQ+P-1} - \mathbf{r}_{nQ+Q-1})(\mathbf{r}_{nQ+P-1} - \mathbf{r}_{nQ+Q-1})^H]$. Under the norm constraint, J_0 is minimized by setting $\mathbf{g}$ as the eigenvector corresponding to the minimum eigenvalue of $\bar{\mathbf{R}}$.

Taking into account the existence of CP, we can rewrite the cost function as

$$J_0(\mathbf{g}) = \mathbf{g}^H\mathbf{H}_{\text{ibi}}\bar{\mathbf{U}}\mathbf{H}_{\text{ibi}}^H\mathbf{g} + 2\sigma_w^2 \tag{17}$$

where $\mathbf{H}_{\text{ibi}}$ consists of the last $L+M-P$ columns of $\mathbf{H}$ and an $(L+M-P)\times(L+M-P)$ matrix $\bar{\mathbf{U}} = E[(\bar{\mathbf{u}}_{nQ-1} - \bar{\mathbf{u}}_{nQ+N-1})(\bar{\mathbf{u}}_{nQ-1} - \bar{\mathbf{u}}_{nQ+N-1})]^H$ where $\bar{\mathbf{u}}_i = \sum_{k=1}^{K}[u^{(k)}[i] \ \cdots \ u^{(k)}[i-(L+M-P)+1]]^T$. When J_0 is minimized, the first term of the right-hand side in (17) becomes zero since $\mathbf{H}_{\text{ibi}}$ has full column rank due to AS2). Then, we have

$$\bar{\mathbf{U}}\mathbf{H}_{\text{ibi}}^H\mathbf{g} = \bar{\mathbf{U}}\mathbf{c}_{\text{ibi}} = \mathbf{0}. \tag{18}$$

where $\mathbf{c}_{\text{ibi}} = \mathbf{H}_{\text{ibi}}^H\mathbf{g}$ represents IBI. If $\bar{\mathbf{U}}$ has full rank, $\mathbf{c}_{\text{ibi}} = \mathbf{0}$. In MC-CDMA systems, the correlation matrix $\bar{\mathbf{U}}$ becomes

$$\bar{\mathbf{U}} = \sum_{k=1}^{K}\left(\mathbf{F}_1^H\mathbf{B}^{(k)}\mathbf{F}_1 + \mathbf{F}_2^H\mathbf{B}^{(k)}\mathbf{F}_2\right) \tag{19}$$

where $\mathbf{B}^{(k)} = E[|d_n^{(k)}|^2\mathbf{s}^{(k)}\mathbf{s}^{(k)H}] = \sigma_d^2\mathbf{s}^{(k)}\mathbf{s}^{(k)H}$ is a rank-one matrix, $\mathbf{F}_1 = [\mathbf{f}_{N-1} \ \cdots \ \mathbf{f}_{N-(L+M-P)}]$, and $\mathbf{F}_2 = [\mathbf{f}_{N-P-1} \ \cdots \ \mathbf{f}_{N-(L+M)}]$, where $\mathbf{f}_l, l = 0, \cdots, N-1$, is the lth column of $\mathbf{F}$. Since $\bar{\mathbf{U}}$ is the sum of $2K$ matrices of rank one, the rank of $\bar{\mathbf{U}}$ is less than or equal to $2K$ [18]. If $2K < L+M-P-1$, $\bar{\mathbf{U}}$ is rank deficient and (18) has nontrivial solutions $\mathbf{c}_{\text{ibi}} \neq \mathbf{0}$. Thus, an immediate necessary condition for $\bar{\mathbf{U}}$ to have full rank is

$$K \geq \frac{1}{2}(L+M-P). \tag{20}$$

This theoretical result suggests that MERRY fails to shorten a channel when the number of users is small. On the other hand, when the number of users is large, $\bar{\mathbf{U}}$ can have full rank depending on the used spreading codes, and then MERRY can shorten a channel properly.

3.2 Channel Shortening Method Based on Uncorrelation Property among Different Blocks

We briefly explain a conventional shortening method presented in [16]. For easy understanding, suppose that the noise can be neglected. In the second-order statistical method in [16], a TEQ vector $\mathbf{g}$ is chosen so that

$$\check{\mathbf{R}}\mathbf{g} = \mathbf{0} \tag{21}$$

where $\check{\mathbf{R}} = E\left[r_0[nQ-1]\check{\mathbf{\Gamma}}_n\right]$ and $\check{\mathbf{\Gamma}}_n$ is a $\Delta' \times LD$ matrix whose ith row is $\mathbf{r}_{nQ+P+i}^H, i = 0, \cdots, \Delta'-1$, where $\Delta' \geq L+M-P-1$. The matrix $\check{\mathbf{\Gamma}}_n$ can be rewritten as

$$\check{\mathbf{\Gamma}}_n = \begin{bmatrix} \mathbf{r}_{nQ+P}^H \\ \vdots \\ \mathbf{r}_{nQ+P+\Delta'-1}^H \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{u}_{nQ+P}^H \\ \vdots \\ \mathbf{u}_{nQ+P+\Delta'-1}^H \end{bmatrix}}_{\check{\mathbf{\Sigma}}_n}\mathbf{H}^H = \left[\check{\mathbf{\Sigma}}_{0,n} \ \check{\mathbf{\Sigma}}_{1,n}\right]\mathbf{H}^H \tag{22}$$

where $\check{\mathbf{\Sigma}}_{0,n}$ consists of the first $P+1$ columns of $\check{\mathbf{\Sigma}}_n$ and $\check{\mathbf{\Sigma}}_{1,n}$ is the remaining part. Note that $\check{\mathbf{\Sigma}}_{0,n}$ does not contain the transmitted samples of the $n-1$th block. Due to an uncorrelation property among different transmitted blocks, we have

$$E\left[r_0[nQ-1]\left[\check{\mathbf{\Sigma}}_{0,n} \ \check{\mathbf{\Sigma}}_{1,n}\right]\right] = \left[\mathbf{0}_{\Delta'\times(P+1)} \ \ \check{\mathbf{\Phi}}\right] \tag{23}$$

where $\check{\mathbf{\Phi}} = E\left[r_0[nQ-1]\check{\mathbf{\Sigma}}_{1,n}\right]$. Finally, we have $\check{\mathbf{\Phi}}\mathbf{c}_{\text{ibi}} = \mathbf{0}$. Even in the cases of correlated sources, IBI can be canceled as long as $\check{\mathbf{\Phi}}$ has full column rank. According to Lemma 5 in [16], it can be shown that $\check{\mathbf{\Phi}}$ has full column rank unless $\phi_{N-1,N-1-m} = 0, m = 0, \cdots, M$, where $\phi_{ij} = E[x_n[i]x_n^*[j]]$ and $x_n[i]$ is the ith entry of $\mathbf{x}_n = \sum_{k=1}^{K}\mathbf{x}_n^{(k)}$.

As explained above, the conventional method (21) can shorten a channel in the ideal situation where the received signal statistics is available. In practice, the ensemble average in $\check{\mathbf{R}}$ is replaced by the time average. It is desirable that satisfactory performance can be obtained using a small number of blocks because of the practical limitation of latency and hardware complexity. Simulations in [16] revealed that this method requires a large number of blocks for time-averaging. This disadvantage comes from the use of only one sample $r_0[nQ-1]$ to obtain $\check{\mathbf{R}}$. When the correlation between $r_0[nQ-1]$ and $\check{\mathbf{\Gamma}}_n$ is weak, a large number of blocks are required to acquire precisely the statistics in (23). To overcome this disadvantage, we propose an improved method in the next section.

4. Improved Channel Shortening Method

4.1 Basic Idea

We begin by defining the TEQ output autocorrelation ρ_δ as

$$\rho_\delta = E[y[nQ-1]y^*[nQ+P+\delta]],\ \delta = 0, \cdots, \Delta-1 \qquad (24)$$

where a time-lag parameter Δ is a positive constant. If a channel $\{c[i]\}$ is shortened, the TEQ output $y[i]$ in the absence of noise can be written as

$$y[i] = \sum_{k=1}^{K}\sum_{l=0}^{P} c[l]u^{(k)}[i-l]. \qquad (25)$$

In this case, $y[nQ-1]$ contains transmitted samples $u^{(k)}[i], i = nQ-P-1, \cdots, nQ-1$, and $y[nQ+P+\delta]$ contains $u^{(k)}[i], i = nQ+\delta, \cdots, nQ+P+\delta$. Since $y[nQ-1]$ and $y[nQ+P+\delta]$ do not share a common transmitted sample, and the transmitted block is uncorrelated among different blocks, the autocorrelation becomes zero $\rho_\delta = 0$ for all δ. This observation motivates us to develop a TEQ design based on ρ_δ.

Now, we propose the following cost function

$$J(\mathbf{g}) = \sum_{\delta=0}^{\Delta-1} \| E\left[\mathbf{r}_{nQ-1}y^*[nQ+P+\delta]\right] \|^2. \qquad (26)$$

Let us define a correlation matrix

$$\mathbf{R}_\delta = E\left[\mathbf{r}_{nQ-1}\mathbf{r}_{nQ+P+\delta}^H\right]. \qquad (27)$$

Since $E[\mathbf{r}_{nQ-1}y^*[nQ+P+\delta]] = \mathbf{R}_\delta\mathbf{g}$, the cost function can be rewritten as

$$J(\mathbf{g}) = \mathbf{g}^H\left(\sum_{\delta=0}^{\Delta-1}\mathbf{R}_\delta^H\mathbf{R}_\delta\right)\mathbf{g} = \mathbf{g}^H\mathcal{R}^H\mathcal{R}\mathbf{g} \qquad (28)$$

where

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}_0 \\ \vdots \\ \mathbf{R}_{\Delta-1} \end{bmatrix}. \qquad (29)$$

Suppose that the minimum of J is zero. Then, if J is minimized, we have $\mathcal{R}\mathbf{g} = \mathbf{0}$. This means $\mathbf{R}_\delta\mathbf{g} = \mathbf{0}$ and thus $\rho_\delta = \mathbf{g}^H\mathbf{R}_\delta\mathbf{g} = 0$ for all δ. What we have to do next is to prove that a channel can be shortened when J is minimized. We prove it in 4.2.

Another straightforward choice of the cost function is

$$C(\mathbf{g}) = \sum_{\delta=0}^{\Delta-1}|\rho_\delta|^2 = \sum_{\delta=0}^{\Delta-1}|\mathbf{g}^H\mathbf{R}_\delta\mathbf{g}|^2. \qquad (30)$$

Clearly, if $C = 0$, we have $\rho_\delta = 0$. Interestingly, this cost function is similar to that of the SAM algorithm [13] except that the autocorrelation is defined as a function of time difference from the timing $i = nQ-1$ in C. However, a difficulty of minimization of C arises from its nonconvexity. Thus, in the following, we focus on J rather than C.

4.2 Analysis of Shortening Capability

Before we show the main result, we need to define some matrices. We rewrite the correlation matrix $\mathbf{R}_\delta$ as $\mathbf{R}_\delta = \mathbf{H}\mathbf{U}_\delta\mathbf{H}^H$ where $\mathbf{U}_\delta = E\left[\mathbf{u}_{nQ-1}\mathbf{u}_{nQ+P+\delta}^H\right]$. Let us define an $(L+M)\Delta \times (L+M)$ matrix $\mathbf{U}$ as

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_0 \\ \vdots \\ \mathbf{U}_{\Delta-1} \end{bmatrix}, \qquad (31)$$

and let $\tilde{\mathbf{U}}$ be an $(L+M)\Delta\times(L+M-P-1)$ matrix consisting of the last $L+M-P-1$ columns of $\mathbf{U}$. Now we can obtain an important result on the shortening capability of the proposed method.

Proposition: A channel $\{c[i]\}$ is shortened to within $P+1$ by minimizing J subject to the norm constraint $\|\mathbf{g}\| = 1$ if the following conditions hold:

· AS1) and AS2).
· Noise is negligible.
· $\tilde{\mathbf{U}}$ has full column rank.

Proof: The minimum of J is the minimum eigenvalue of $\mathcal{R}^H\mathcal{R}$. Since $\mathbf{H}$ is tall, the minimum eigenvalue of $\mathcal{R}^H\mathcal{R}$ is zero. Thus, it is sufficient to prove that $c[i] = 0, i = P+1, \cdots, L+M-1$, when $\mathcal{R}\mathbf{g} = \mathbf{0}$ holds. The vectors $\mathbf{u}_{nQ-1}$ and $\mathbf{u}_{nQ+P+\delta}$ are given by

$$\mathbf{u}_{nQ-1} = \sum_{k=1}^{K}\begin{bmatrix} u^{(k)}[nQ-1] \\ \vdots \\ u^{(k)}[nQ-(L+M)] \end{bmatrix}, \qquad (32)$$

$$\mathbf{u}_{nQ+P+\delta} = \sum_{k=1}^{K}\begin{bmatrix} u^{(k)}[nQ+P+\delta] \\ \vdots \\ u^{(k)}[nQ] \\ u^{(k)}[nQ-1] \\ \vdots \\ u^{(k)}[nQ+P+\delta-(L+M)-1] \end{bmatrix}, \qquad (33)$$

respectively. All entries contained in $\mathbf{u}_{nQ-1}$ come from the $n-1$th block. Meanwhile, the top $P+\delta+1$ entries of $\mathbf{u}_{nQ+P+\delta}$

come from the nth block and the remaining entries from $n-1$th block. Owing to AS1), all entries in the first $P+\delta+1$ columns of $\mathbf{U}_\delta$ become zero. Thus, $\mathbf{U}$ takes the form

$$\mathbf{U} = [\underbrace{\mathbf{0}}_{P+1} \quad \underbrace{\tilde{\mathbf{U}}}_{L+M-P-1}] \tag{34}$$

The matrix $\mathcal{R}$ can be written as

$$\mathcal{R} = \underbrace{\begin{bmatrix} \mathbf{H} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{H} \end{bmatrix}}_{\tilde{\mathbf{H}}} \mathbf{U}\mathbf{H}^H = \tilde{\mathbf{H}} \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{U}} \end{bmatrix} \mathbf{H}^H. \tag{35}$$

Now, we consider $\mathcal{R}\mathbf{g} = \mathbf{0}$. Because $\mathbf{H}$ is assumed to have full column rank due to AS2), $\tilde{\mathbf{H}}$ has also full column rank and we get $\tilde{\mathbf{U}}\mathbf{H}_{\text{ibi}}^H\mathbf{g} = \mathbf{0}$. Furthermore, since $\tilde{\mathbf{U}}$ is assumed to have full column rank, we have $\mathbf{H}_{\text{ibi}}^H\mathbf{g} = \mathbf{0}$. This implies $c[i] = 0, i = P+1, \cdots, L+M-1$. □

Let us look closely at the matrix $\tilde{\mathbf{U}}$. The nth transmitted block of the kth user can be represented as

$$\begin{bmatrix} u^{(k)}[nQ] \\ \vdots \\ u^{(k)}[nQ+P-1] \\ u^{(k)}[nQ+P] \\ \vdots \\ u^{(k)}[nQ+Q-1] \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{N-P}^H\mathbf{s}^{(k)} \\ \vdots \\ \mathbf{f}_{N-1}^H\mathbf{s}^{(k)} \\ \mathbf{f}_0^H\mathbf{s}^{(k)} \\ \vdots \\ \mathbf{f}_{N-1}^H\mathbf{s}^{(k)} \end{bmatrix} d_n^{(k)} \tag{36}$$

where $\mathbf{f}_i^H$ is the ith row of $\mathbf{F}$. Then, the matrix $\mathbf{U}_\delta$ can be written as

$$\mathbf{U}_\delta = \sum_{k=1}^{K} \begin{bmatrix} \underbrace{\mathbf{0}}_{P+1+\delta} & \sigma_{N-1,N-1}^{(k)} & \cdots \sigma_{N-1,N-(L+M-P-\delta-1)}^{(k)} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \sigma_{N-(L+M),N-1}^{(k)} \cdots & \sigma_{N-1,N-1}^{(k)} \end{bmatrix} \tag{37}$$

where $\sigma_{i,j}^{(k)} = \left(\mathbf{f}_i^H\mathbf{s}^{(k)}\right)\left(\mathbf{f}_j^H\mathbf{s}^{(k)}\right)^*$. Equation (37) implies that $\mathbf{U}_\delta$ depends only on the used spreading codes. For example, let us consider a special case when $K=1$. If $\Delta = L+M-P-1$ and

$$\sigma_{N-i,N-1}^{(1)} \neq 0 \tag{38}$$

for any $i, 1 \leq i \leq L+M$, it is clear that $\tilde{\mathbf{U}}$ has full column rank. Thus, the rank of $\tilde{\mathbf{U}}$ depends on the used spreading codes. Note that Δ also influences on the rank of $\tilde{\mathbf{U}}$. Therefore, when the spreading codes and Δ are given, we can easily check if $\tilde{\mathbf{U}}$ has full column rank. The condition (38) can be satisfied by some Walsh codes.

The proposed method is similar to the conventional method described in 3.2 since both methods exploit the uncorrelation property among different transmitted blocks. The main difference between these two methods is that, in the proposed method, the correlation matrix $\mathbf{R}_\delta$ is obtained by using LD samples in $\mathbf{r}_{nQ-1}$ rather than only one sample $r_0[nQ-1]$. We expect that the increase of the number of samples leads to the reduction of required blocks for time-averaging. The validity of this idea is evaluated through simulation in Sec. 5.

4.3 Implementation

The proposed method consists of the following steps:

Method1:

Step1) Estimate $\mathbf{R}_\delta$ for $\delta = 0, \cdots, \Delta-1$ by time-averaging over B_0 blocks:

$$\hat{\mathbf{R}}_\delta = \frac{1}{B_0}\sum_{n=1}^{B_0} \mathbf{r}_{nQ-1}\mathbf{r}_{nQ+P+\delta}^H. \tag{39}$$

Step2) According to (29), construct the matrix $\hat{\mathcal{R}}$ using $\hat{\mathbf{R}}_\delta$.

Step3) Set $\mathbf{g}$ as the eigenvector corresponding to the minimum eigenvalue of $\hat{\mathcal{R}}^H\hat{\mathcal{R}}$.

In our preliminary simulations, we confirmed that a channel can be shortened by the above procedure. As shown in the next section; however, we found that its symbol error rate performance is unsatisfactory. The reason is two-fold. First, the magnitude of the effective channel $||\mathbf{c}||$ becomes small and thus the desired component is weakened. Second, the frequency response of a TEQ is not flat, and this leads to noise enhancement. In order to overcome these problems, we put the following constraints:

$$E[|y[nQ-1]|^2] = 1, \tag{40}$$
$$|\tilde{\mathbf{f}}_i\mathbf{g}_d|^2 - |\tilde{\mathbf{f}}_j\mathbf{g}_d|^2 < \varepsilon, \tag{41}$$

for $d = 0, \cdots, D-1, i = 0, \cdots, N-1, j > i$, where $\mathbf{g}_d = [g_d[0] \ \cdots \ g_d[L-1]]^T$, $\tilde{\mathbf{f}}_i$ is the $1 \times L$ vector consisting of the first L entries of the ith row of $\mathbf{F}$, ε is a small positive constant that controls the flatness of the frequency response of a TEQ. The first constraint prevents to weaken the desired component in $\mathbf{c}$. Due to the second constraint, the frequency gain of a TEQ approaches flat. Eventually, the modified steps are as follows:

Method2:

Step1) and **Step2)** are the same as in **Method1**.

Step3) Estimate an autocorrelation matrix by time-averaging over B_0 blocks:

$$\hat{\mathbf{R}} = \frac{1}{B_0}\sum_{n=1}^{B_0} \mathbf{r}_{nQ-1}\mathbf{r}_{nQ-1}^H. \tag{42}$$

Step4) Solve the following quadratically constrained quadratic programming (QCQP) problem:

$$\begin{aligned} &\min \ \ \mathbf{g}^H\hat{\mathcal{R}}^H\hat{\mathcal{R}}\mathbf{g} \\ &\text{sub. to} \ \ \mathbf{g}^H\hat{\mathbf{R}}\mathbf{g} = 1 \\ &\qquad \mathbf{g}^H\left(\mathbf{P}_d^T\tilde{\mathbf{f}}_i^H\tilde{\mathbf{f}}_i\mathbf{P}_d - \mathbf{P}_d^T\tilde{\mathbf{f}}_j^H\tilde{\mathbf{f}}_j\mathbf{P}_d\right)\mathbf{g} < \epsilon, \\ &\qquad d = 0, \cdots, D-1, i = 0, \cdots, N-1, j > i \end{aligned} \tag{43}$$

where $\mathbf{P}_d$ is an $L \times LD$ matrix whose lth row is the $lD+d$th row of $LD \times LD$ identity matrix. The above QCQP problem

[19] can be solved using well-known solvers SeDuMi [20] and YALMIP [21]. We should note that, the norm constraint does not necessarily hold in Method2. As shown in the next section, however, Method2 can provide a significant performance improvement.

5. Simulation Results

We conducted simulations to evaluate the performance of the proposed method in the presence of noise, and compared it with a conventional method in [16], the MERRY algorithm [12], and the receiver without TEQ. Unless otherwise stated, we used the parameters listed in Table 1. A channel coefficient $h_d[i]$ has the variance $\sigma_m^2 = \lambda \exp(-\alpha d), d = 0, \cdots, M$, where $\alpha = 0.1$ and λ ensures the unit average energy. Symbol error rate (SER) is obtained by averaging over B_1 blocks and B_2 trials (Each trial has a different channel). All receivers use zero-forcing FEQs built from the exact knowledge of the shortened channel after performing channel shortening. The average received SNR was defined by

$$\mathrm{SNR} = \frac{1}{DKQ\sigma_w^2} \sum_{i=0}^{Q-1} \sum_{d=0}^{D-1} E[|r_d[nQ+i]|^2]. \tag{44}$$

Though the FFT size N was set to small to save computation time, the obtained results are enough to validate our theoretical results and clarify the differences among various methods.

First, Fig. 2 shows an example of the impulse response of the original physical channel and the effective channels obtained by the proposed method, the conventional method, and MERRY. We can observe that the proposed method successfully shortens a channel to within $P + 1$.

Next, we consider the effect of channel noise. Figure 3 illustrates symbol error rate (SER) for various SNR. From this figure, it can be observed that Method1 has unsatisfactory performance, while Method2 outperforms the other methods. SER performance in the absence of IBI (no IBI) is also depicted in this figure. This curve can be regarded as a performance limit. In low SNR region, there is little room for improvement. Actually, the performance of Method2 is almost the same as that of the receiver without TEQ. In terms of practical use, it should be emphasized that Method2 provides a significant performance improvement in high SNR region. Though the performance of Method2 approaches the performance limit, there is still a performance gap between them. The reason of this gap is that the number of blocks used to estimate correlation matrices is insufficient. Though not shown, we have confirmed that this gap can be reduced using a larger number of blocks.

Next, SER performances when the number of users varies is illustrated in Fig. 4, where we set SNR is 30 dB, relatively high. The performance of the proposed method is almost the same as that of the conventional method when the number of users is very small. However, the proposed method is superior to the conventional one as the number of users increases. Interestingly, the performance of MERRY becomes better when the number of users is large. This is because, as mentioned in 3.1, the matrix $\bar{\mathbf{U}}$ can have full rank when the number of users increases.

Next, we show how the number of blocks B_0 used to estimate correlation matrices effects the performance. Figure 5 shows SER as a function of B_0, where SNR=30 dB. We can observe from this result that the proposed method requires a less number of blocks for a target SER. For exam-

Table 1 Simulation specifications.

The number of subcarriers N	16
CP length P	4
The order of physical channels M	7
TEQ filter length L	8
The number of users K	4
The number of antennas D	2
Modulation scheme	QPSK
Spreading codes	Walsh-Hadamard sequences
Gain flatness parameter ε	0.01
Time-lag parameter Δ, Δ'	$L + M - P - 1 = 10$
The number of blocks used to estimate correlation matrices B_0	10^3
The number of blocks used to evaluate errors for each trial B_1	10^3
The number of trials B_2	10^4

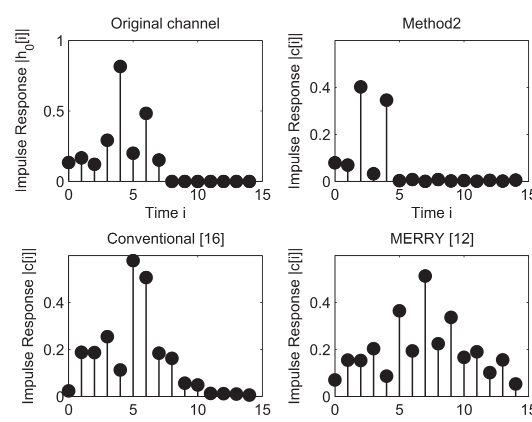

Fig. 2 Impulse response magnitude.

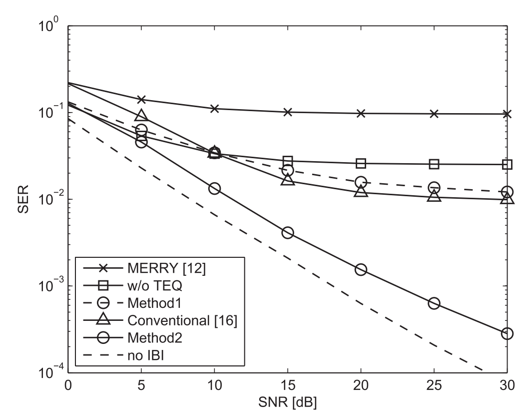

Fig. 3 SER versus SNR.

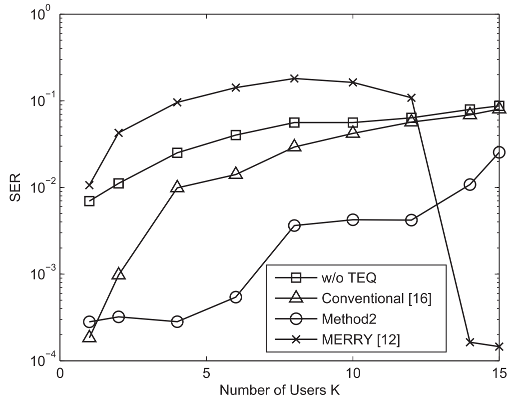

Fig. 4 SER versus the number of users.

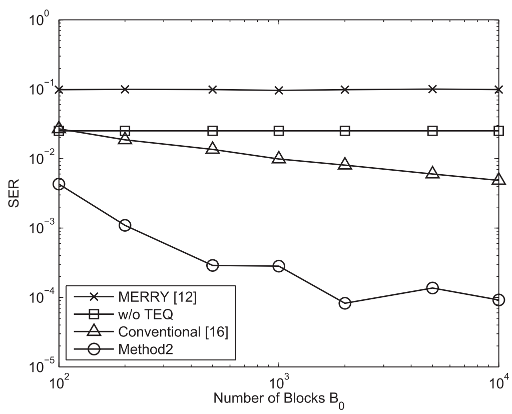

Fig. 5 SER versus the number of blocks used for matrix estimation.

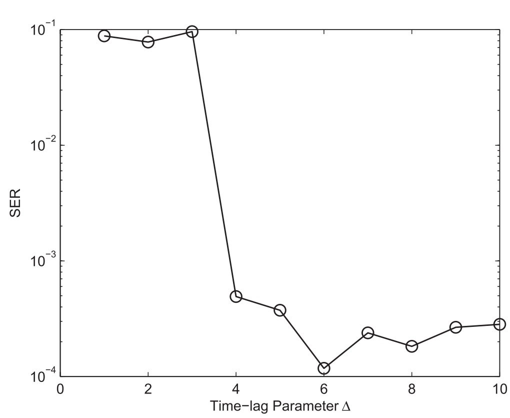

Fig. 6 SER versus time-lag parameter.

ple, to achieve SER of 10^{-2}, the conventional method [16] requires about 1,000 blocks, while the proposed method requires less than 100 blocks. The performance of MERRY does not improve even if B_0 increases because the matrix $\bar{\mathbf{U}}$ is rank-deficient.

Next, we consider the effect of the time-lag parameter Δ. Figure 6 illustrates the SER when Δ varies, where SNR is 30 dB. The performance is almost the same if Δ is more than 4. Though Δ should be $L + M - P - 1 = 10$ in the case of $K = 1$ as mentioned in 4.2, Δ can be smaller than

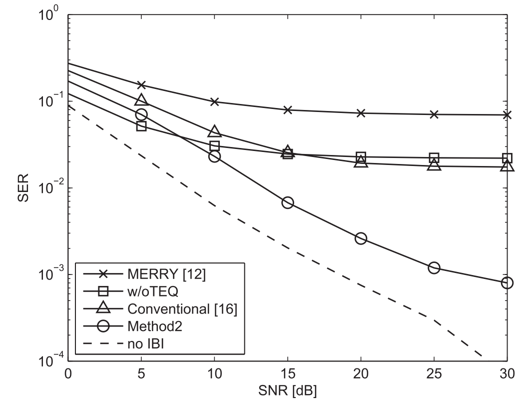

Fig. 7 SER versus SNR when orthogonal M-sequences are used.

$L + M - P - 1$ when $K = 6$.

Finally, we discuss the effect of the spreading codes. Figure 7 shows SER for various SNR using the orthogonal M-sequences [22]. The performance of the proposed method is almost the same as in the case of Walsh-Hadamard sequences. It can be observed that Method2 has SER floor. This is because the number of blocks B_0 is insufficient. This suggests that a more data efficient method is desirable.

6. Conclusions

In this paper, we have proposed a blind channel shortening method, and applied to downlink MC-CDMA systems. We proved that the proposed method can shorten a channel and thus to suppress IBI under certain conditions. By simulations, we have shown that the proposed method requires a less number of received blocks than a conventional method in [16], and provides significant performance improvement comparing to MERRY [12], when the number of users is moderate. The development of a more data efficient method remains a challenging issue.

References

[1] Z. Wang and G.B. Giannakis, “Wireless multicarrier communications,” IEEE Signal Process. Mag., vol.17, no.3, pp.29–48, May 2000.

[2] IEEE Std. 802.16e 2005, “Air interface for fixed and mobile broadband wireless access systems,” Feb. 2006.

[3] S. Hara and R. Prasad, “Overview of multicarrier CDMA,” IEEE Commun. Mag., vol.35, no.12, pp.126–133, Dec. 1997.

[4] F. Adachi, D. Garg, S. Takaoka, and K. Takeda, “Broadband CDMA techniques,” IEEE Wireless Commun., vol.12, no.2, pp.8–18, April 2005.

[5] M. Juntti, M. Vehkapera, J. Leinonen, Z. Li, D. Tujkovic, S. Tsumura, and S. Hara, “MIMO MC-CDMA communications for future cellular systems,” IEEE Commun. Mag., vol.43, no.2, pp.118–124, Feb. 2005.

[6] S. Plass and S. Kaiser, “MC-CDMA versus OFDMA in cellular environments,” Proc. EUSIPCO, 2005.

[7] A. Phasouliotis and D.K.C. So, “Performance analysis and comparison of downlink MIMO MC-CDMA and MIMO OFDMA systems,” Proc. IEEE 69th VTC Spring, pp.1–6, April 2009.

[8] D. Kim and G.L. Stuber, "Residual ISI cancellation for OFDM with applications to HDTV broadcasting," IEEE J. Sel. Areas Commun., vol.50, no.12, pp.2136–2148, Dec. 2002.
[9] S. Hori, N. Kikiuma, and N. Inagaki, "MMSE adaptive array suppressing only multipath waves with delay times beyond the guard interval for fixed reception in the OFDM systems," IEICE Trans. Commun. (Japanese Edition), vol.J86-B, no.9, pp.1934–1940, Sept. 2003.
[10] P.J.W. Melsa, R.C. Younce, and C.E. Rohrs, "Impulse response shortening for discrete multitone transceivers," IEEE Trans. Commun., vol.44, no.12, pp.1662–1672, Dec. 1996.
[11] R.K. Martin and C.R. Johnson, Jr., "Adaptive equalization: Transitioning from single-carrier to multicarrier systems," IEEE Signal Process. Mag., vol.22, no.6, pp.108–122, Nov. 2005.
[12] R.K. Maritn, J. Balakrishnan, W.A. Sethares, and C.R. Johnson, Jr., "A blind, adaptive TEQ for multicarrier systems," IEEE Trans. Signal Process., vol.9, no.11, pp.341–343, Nov. 2002.
[13] J. Balakirishnan, R.K. Martin, and C.R. Johnson, Jr., "Blind, adaptive channel shortening by sum-squared auto-correlation minimization (SAM)," IEEE Trans. Signal Process., vol.51, no.12, pp.3086–3093, Dec. 2003.
[14] R. Romano and S. Barbarossa, "Non-data aided adaptive channel shortening for efficient multi-carrier systems," Proc. IEEE ICASSP, vol.4, pp.233–236, Hong Kong, April 2003.
[15] T. Miyajima and Z. Ding, "Second-order statistical approaches to channel-shortening in multicarrier systems," IEEE Trans. Signal Process., vol.52, no.11, pp.3253–3264, Nov. 2004.
[16] T. Miyajima and Y. Watanabe, "Blind channel shortening for block transmission of correlated signals," IEICE Trans. Fundamentals, vol.E91-A, no.11 pp.3095–3103, Nov. 2008.
[17] T. Miyajima and Z. Ding, "Subcarrier nulling algorithms for channel shortening in uplink OFDMA systems," IEEE Trans. Signal Process., vol.60, no.5, pp.2374–2385, May 2012.
[18] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, 1985.
[19] Z.-Q. Luo, W.-K. Ma, A.M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," IEEE Signal Process. Mag., vol.27, no.3, pp.20–34, May 2010.
[20] J.F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," Optimization Methods and Software, vol.11, no.1–4, pp.625–653, 1999, http://sedumi.ie.lehigh.edu, accessed Feb. 2. 2012.
[21] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," Proc. IEEE CASCSD, pp.284–289, Taipei, Sept. 2004, http://users.isy.liu.se/johanl/yalmip, accessed Feb. 2. 2012.
[22] H. Habuchi, "Pseudo-noise sequences based on M-sequence and its application for communications," IEICE Fundamentals Review, vol.3, no.1, pp.32–42, July 2009 (in Japanese).

Mizuki Kotake received B.Eng. degree in electrical and electronic engineering from Ibaraki University, Hitachi, Japan, in 2011. He is currently a master course student at Ibaraki University. His current interests are in signal processing for communications.

Teruyuki Miyajima received B.Eng., M.Eng., and Ph.D. degrees in electrical engineering from Saitama University, Saitama, Japan, in 1989, 1991, and 1994, respectively. In 1994, he joined Ibaraki University, Hitachi, Japan, where he is currently an associate professor in the Department of Electrical and Electronic Engineering. His current interests are in statistical signal processing for communications. Dr. Miyajima is a member of IEEE.