<table>
<thead>
<tr>
<th>Title</th>
<th>Spectrally Efficient Frame Format-Aided Turbo Equalization with Channel Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Takano, Yasuhiro; Anwar, Khoirul; Matsumoto, Tad</td>
</tr>
<tr>
<td>Citation</td>
<td>IEEE Transactions on Vehicular Technology, 62(4): 1635-1645</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-11-30</td>
</tr>
<tr>
<td>Type</td>
<td>Journal Article</td>
</tr>
<tr>
<td>Text version</td>
<td>author</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10119/10859">http://hdl.handle.net/10119/10859</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Copyright (C) 2012 IEEE. IEEE Transactions on Vehicular Technology, 62(4), 2012, 1635-1645. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.</td>
</tr>
</tbody>
</table>

Description

This is the author's version of the work.
Abstract—Chained turbo equalization (CHATUE) has been recently recognized as a low-complexity frequency domain turbo equalization technique that eliminates the necessity of transmitting the cyclic prefix (CP), and hence allows for spectrally efficient signalling in wireless communications. However, two issues arise from the original version of CHATUE (referred to as CHATUE1) as a consequence of eliminating the CP, which are the noise enhancement and the latency due to the time-concatenated structure. This paper proposes a new version of CHATUE (referred to as CHATUE2) to solve the noise enhancement problem. CHATUE2 retrieves the circulant structure of the channel matrix, originally inherent within the CP-transmission, by utilizing composite replica signals that combines the received and the soft reference signals replicated from the log-likelihood ratio fed back from the decoder. For this purpose, this paper determines the optimal combining ratio based on the minimum mean-square-error criterion. CHATUE2 is hence able to achieve an improvement in bit-error-rate (BER) performance over CHATUE1. In addition, this paper provides a solution to solving the latency problem by making a practical assumption on the training sequence (TR) transmission which is required to perform channel estimation generally, in practical systems. Furthermore, this paper proposes a new channel estimation technique, chained turbo estimation (CHATES), which improves the spectrum efficiency and asymptotically achieves the Cramér-Rao bound. CHATES assumes that the TR length is exactly equal to the channel impulse response length, although the conventional technique requires twice as long as or even longer TR lengths. Numerical results show that CHATUE2 with CHATES achieves 1 dB gain over conventional turbo equalization with a CP at $10^{-5}$ BER in realistic propagation scenarios represented by channel sounding measurement data as well as in model-based frequency-selective fading channels.

Index Terms—chained turbo equalization (CHATUE), chained turbo estimation (CHATES), cyclic prefix, spectrum efficiency, transmission frame format.

I. INTRODUCTION

Cyclic prefix (CP) aided block transmission has been recently gaining popularity in block transmission systems such as in single carrier frequency division multiple access (SC-FDMA) and/or orthogonal frequency division multiple access (OFDMA). One of the benefits of utilizing CP is to reduce the computational complexity for signal detection while keeping the robustness against fading frequency selectivity. The CP-transmission, on the other hand, imposes an overhead in the transmission format structure. It is hence preferable to minimize the length of the CP to improve the transmission energy- and spectrum-efficiencies. However, it causes serious degradation in bit-error-rate (BER) performance if the length of the CP is shorter than the actual length of the channel impulse response (CIR). Chained turbo equalization (CHATUE) proposed in [1] provides a solution to this problem: CHATUE makes it possible to perform the frequency domain equalization processing, even without a CP, while requiring the same order of computational complexity as that of conventional frequency domain turbo equalization with CP transmission (TEQ-CP) [2]. Since CHATUE requires no CP-transmission, it provides us with more design flexibility in terms of energy- and spectral-efficiency tradeoff. In other words, CHATUE enables us to transmit more information bits or to use a lower rate code by utilizing the time duration allocated for a CP. Thereby, CHATUE has a potential to improve performance over TEQ-CP, as detailed in [3], in terms of required signal-to-noise power ratio (SNR) or throughput efficiency. Nevertheless, the original CHATUE (referred to as CHATUE1) has the following two problems, which are the consequence of eliminating CP-transmission:

1) Latency: CHATUE algorithms studied so far in [1], [3], [4] require a processing latency three times that of TEQ-CP, since it performs iterations over at least three blocks (past, current and future blocks) to cancel the inter-block-interference (IBI). On the other hand, TEQ-CP performs turbo iterations within the current block alone.

2) Noise Enhancement: CHATUE1 utilizes a so called J-matrix [5] to retrieve the circulant structure of the channel matrix. However, a part of the signal after the transformation suffers from noise enhancement because of the multiplication of the J-matrix, as detailed in Section III-D. The SNR at the output of equalization with CHATUE1, as a consequence, is decreased compared to that of TEQ-CP.

This paper shows that Problem 1) can be easily solved under a practical assumption on the training sequence transmission. For Problem 2), this paper proposes a novel algorithm, CHATUE version 2 (referred to as CHATUE2).

Furthermore, this paper proposes a new channel estimation technique, chained turbo estimation (CHATES), that inherits the CHATUE concept, to pursue further improvement of the spectrum efficiency. The required length $N_t$ of the training sequence (TR) is determined according to the length $W$ of CIR. Conventional least-squares-based techniques requires
$N_t \geq 2W$ to achieve accurate channel estimates if the transmission format does not have a guard interval (GI) between the TR and its neighboring segments. However, CHATES requires a TR length of only $N_t = W$, while it achieves the Cramér-Rao bound (CRB) asymptotically.

In this paper, the performance of the proposed techniques are verified through computer simulations in realistic propagation scenarios represented by channel sounding measurement data as well as in model-based scenarios.

A. Organization of this paper

This paper is organized as follows. Section II describes the system model assumed in this paper. Section III reviews CHATUE1, discusses the above-mentioned problems 1) and 2) in detail, and proposes CHATUE2. Section IV proposes the new turbo channel estimation technique, CHATURE. Section V presents results of computer simulations conducted to verify the effectiveness of the proposed techniques. This paper is concluded in Section VI with some concluding remarks.

B. Notations

The bold upper-case $X$ and lower-case $x$ denote a matrix and a vector, respectively. $X^H$ denotes the transposed conjugation of the matrix $X$, while $\text{diag}(X)$ forms a diagonal matrix from the diagonal elements of its argument matrix $X$. svd($X$) = $UDV^H$ is the singular value decomposition of a matrix $X \in \mathbb{C}^{M \times N}$, where $U \in \mathbb{C}^{M \times M}$ and $V \in \mathbb{C}^{N \times N}$ are unitary matrices and $D \in \mathbb{C}^{M \times N}$ is a rectangular diagonal matrix. $X_{i|j}$ is a submatrix of the first $r$ column vectors in a matrix $X$. Similarly, $x_{i|j}$ is a subvector of the original vector $x$ which extracts the $i$-th to the $j$-th elements from the vector $x$.

II. SYSTEM MODEL

The system model assumed in this paper is depicted in Fig. 1. A binary data information sequence $b(i)$, $1 \leq i \leq N_B N_d R_e$, is encoded by a rate $R_e$ convolutional code (CC) with generator polynomials $(g_1, \ldots, g_{R_e})$ and is interleaved by an interleaver ($\mathcal{I}$). The interleaved coded frame $c_M(k)$, $1 \leq k \leq N_B N_d$, is divided into $N_B$ bursts such that fading is assumed to be static over each burst. The transmitter transmits $N_d$ binary phase shift keyed (BPSK) symbols $x(k_s;l)$ together with a length $N_t$ symbol training sequence and CP, using single carrier signalling, where $l$ and $k_s$ denote the burst index and the symbol index in a burst, respectively. Fig. 2 shows the structure of the transmission formats assumed in this paper. Note that the number $N_{CP}$ of CP symbols is set at zero with the CHATUE algorithms. The length $N_{GI}$ and $N_{G2}$ guard intervals, following the TR and Data part respectively, are also set at zero when we aim at improving the spectrum efficiency.

The receiver receives the signal $y(k_s;l)$ suffering from intersymbol-interference (ISI) due to fading frequency selectivity.

1For simplicity of the system model, we assume binary modulation in this paper. However, extension to higher order modulation is straightforward.

as well as complex additive white Gaussian noise (AWGN). The maximum ISI length is $L = W - 1$ symbols under the assumption that the CIR length is $W$. The received signal corresponding to the transmitted signal in the current burst $l$ can be described in vector form $y(l) \in \mathbb{C}^{K+L}$, as

$$y(l) = H(l)x(l) + H'(l-1)x'(l-1) + H''(l+1)x''(l+1) + n,$$  \hspace{1cm} (1)

where $x(l)$, $x'(l-1)$ and $x''(l+1)$ denote symbol vectors, respectively, transmitted in the current, past and future burst timings, each of which has $K = N_t + N_{G1} + N_{CP} + N_d + N_{G2}$ symbols. $H(l) \in \mathbb{C}^{(K+L) \times K}$ is a Toeplitz matrix representing the convolution of the transmitted data symbols with the CIR in the current block, with

$$H(l) = \begin{bmatrix} h(1;l) & \cdots & h(1;l) \\ \vdots & \ddots & \vdots \\ h(W;l) & \cdots & h(1;l) \\ \vdots & \ddots & \vdots \\ h(W;l) \\ \end{bmatrix}.$$  \hspace{1cm} (2)

On the other hand, the $(K + L) \times K$ matrices $H'(l-1)$ and $H''(l+1)$ are, respectively, given as follows.

$$H'(l-1) = \begin{bmatrix} O_{L \times (K-L)} & H_r(l-1) \\ O_{K \times (K-L)} & O_{K \times L} \end{bmatrix}$$  \hspace{1cm} (3)

with

$$H_r(l) = \begin{bmatrix} h(W;l) & h(W-1;l) & \cdots & h(2;l) \\ h(W;l) & \cdots & h(3;l) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h(W;l) \end{bmatrix},$$

and

$$H''(l+1) = \begin{bmatrix} O_{K \times L} & O_{K \times (K-L)} \\ H_r''(l+1) & O_{L \times (K-L)} \end{bmatrix}$$  \hspace{1cm} (4)

with

$$H_r''(l) = \begin{bmatrix} h(1;l) \\ h(2;l) \\ \vdots \\ h(L;l) & h(L-1;l) & \cdots & h(1;l) \end{bmatrix}.$$
the transmitted symbols. $CC^{-1}$ outputs the estimates of the transmitted sequence $\hat{b}$ by making a hard decision on the decoder’s \textit{a posteriori} LLR $\lambda_{DEC}^p$ corresponding to $b(i)$ after several iterations. $EST$ and $EQU$ utilize the soft replica of the transmitted symbols $\hat{x}_{EST}$ and $\hat{x}_{EQU}$, respectively. $\hat{x}_{d,EST}$ is generated from the \textit{a priori} LLR $\lambda_{EST}^a$ for channel estimation after interleaving the \textit{a posteriori} LLR $\lambda_{DEC}^p$. On the other hand, $\hat{x}_{d,EQU}$ is generated from the equalizer’s \textit{a priori} LLR $\lambda_{EQU}^a$ which is the interleaved version of the extrinsic LLR $\lambda_{DEC}^e = \lambda_{DEC}^p - \lambda_{DEC}^a$, according to the turbo principle.

### III. CHATUE ALGORITHMS

This section reviews CHATUE1 and identifies the causes of the two problems described in Section I, and then proposes a new version of CHATUE, CHATUE2. First of all, the latency issue, raised as Problem 1), is discussed in the framework of the CHATUE algorithm. We observe that Problem 1) can be solved by adopting a practical and reasonable assumption on the transmission format structure. The signal model is then defined under the assumption and the CHATUE1 algorithm is reviewed. The new algorithm CHATUE2 is proposed in Section III-E, following the analysis of Problem 2).

#### A. CHATUE Algorithms and Latency

The time duration which were used for the CP-transmission in TEQ-CP can be eliminated by the CHATUE algorithms. CHATUE algorithms hence can enhance either spectrum- or energy-efficiencies of the system by transmitting more information bits or utilizing lower rate codes, respectively. As a consequence of eliminating CP-transmission, the current data block suffers from IBI due to the neighboring blocks (past and future) as we can observe from Fig. 2 when $N_i = N_{CP} = N_{G1} = N_{G2} = 0$. CHATUE algorithms can cancel the IBI in the current block, by exchanging the LLRs of transmitted bits between the current and the neighboring blocks. However, in exchange for the spectrum or energy efficiency gains, CHATUE algorithms require a latency of at least three times the conventional TEQ-CP technique because in addition to the iterations to detect the current block, it also has to perform iterations with past and future frames.

Nevertheless, the latency can be avoided, or at least reduced by introducing the following assumption, which are of practical importance and hence very reasonable: the TR is transmitted together with the data part for channel estimation, and the length of the TR is designed to be longer than the maximum length of the CIR. As we can observe in Fig. 2, when $N_i > L$ and $N_{CP} = N_{G1} = N_{G2} = 0$, IBI’s occurring in a data segment in the current burst are caused by TRs in the current or future burst. Since the TR pattern is known to the receiver, it is not necessary to exchange the LLR of the transmitted bits between the current and future blocks to detect the IBI symbols.

For the complete detection of the current burst, channel estimation still has to be conducted for the future burst as well as for the current burst. However, it increases the latency slightly since we can estimate the CIR with TR alone. Thereby, the CHATUE algorithms can avoid the necessity of performing iterations over the bursts neighboring in time. CHATUE algorithms only require the latency equivalent to TEQ-CP under an assumption that a TR is transmitted in every burst.

---

**Fig. 1.** System model: Structure of transmitter and receiver.

**Fig. 2.** Structure of burst formats and illustration of input data range for channel estimation.

---

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Alphanumeric</th>
<th>Capitalization</th>
<th>Punctuation</th>
<th>Special Characters</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CC$</td>
<td>$CC^{-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b(i)$</td>
<td>$c(i)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k=\Pi(i)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$TR$</td>
<td>$EST$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EQU$</td>
<td>$\lambda_{EQU}^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{x}_{EST}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{DEC}^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{x}_{EQU}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{DEC}^p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{x}_{d,EST}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{EST}^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{x}_{d,EQU}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{EQU}^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_{EST}^i$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_{EQU}^i$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_i$</td>
<td>$N_{CP}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{G1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{G2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>$\lambda_{DEC}^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This paper assumes each burst is headed by a TR. However, the algorithms described in Section III-E and IV-B can be derived similarly if a TR is allocated at the tail of the burst.

B. Signal Model for CHATUE algorithms with TR transmission

Assuming a TR is transmitted at the head of every burst, we re-formulate the signal model of the CHATUE algorithms. Similarly to (1), the received data segment $y_d(l) \in \mathbb{C}^{N_d+L}$ for the transmitted burst in the current burst timing $l$ is described as

$$y_d(l) = H_d(l)s_d(l) + H_d'(l)s'_d(l) + n_d,$$

where the signal vectors, each of which the size is $N_d \times 1$, are defined as $s_d(l) = x_d(l)$, $s'_d(l) = [0_1 \times (N_d-W) \ x_l(t)]^T$ and $s''_d(l+1) = [x_l(l+1)^T \ 0_1 \times (N_d-W)]^T$. The $\{N_d+L\} \times N_d$ Toeplitz matrices $H_d(l)$, $H'_d(l)$ and $H''_d(l+1)$ are defined in the same way as (2), (3) and (4), utilizing CIR vectors $h(l)$, $h(l)$ and $h(l+1)$, respectively.\(^2\) The $N_d \times 1$ noise vector $n_d$ follows $\mathcal{CN}(0, \sigma^2_n)$.

C. Review of CHATUE version 1 (CHATUE1)

According to [4], the equalizer output of CHATUE1 is given by

$$z_1(l) = \left( I_{N_d} + \Gamma(l) \hat{S}(l) \right)^{-1} \cdot \left[ \Gamma(l)s_d(l) + F^H \hat{\Phi}(l) \Omega(l)^{-1} F \hat{r}_d(l) \right],$$

where $\hat{S}(l) = \text{Diag} \|s_d(l)\|^2$ and

$$\hat{\Phi}(l) = F \hat{H}(l) F^H$$

is a diagonal matrix. The $J$-matrix proposed in [5] are defined as

$$J = \left( \begin{array}{c|c} 0_{(N_d-L) \times L} & I_L \\ \hline I_L & I_{N_d} \end{array} \right) \in \mathbb{R}^{N_d \times (N_d+L)}.$$

$F \in \mathbb{C}^{N_d \times N_d}$ is the DFT matrix whose $(r+1, c+1)$-element is defined as

$$\exp \left[ -2\pi r c \sqrt{-1}/N_d \right] / \sqrt{N_d}$$

with integer indexes $0 \leq r, c \leq N_d - 1$. The matrix $\Omega(l)$ in (6) is given by

$$\Omega(l) = F \Sigma(l) F^H,$$

where

$$\Sigma(l) = \hat{J} \hat{H}(l) \Lambda(l) (\hat{J} \hat{H}(l))^H$$

and

$$\Lambda(l) = E \left[ (\hat{s}_d(l) - s_d(l)) (\hat{s}_d(l) - s_d(l))^H \right],$$

$$\Lambda'(l) = E \left[ (\hat{s}'_d(l) - s'_d(l)) (\hat{s}'_d(l) - s'_d(l))^H \right]$$

and

$$\Lambda''(l+1) = E \left[ (\hat{s}''_d(l+1) - s''_d(l+1)) (\hat{s}''_d(l+1) - s''_d(l+1))^H \right].$$

However, taking into account $\Lambda'(l) = \Lambda''(l+1) = 0$, because $\hat{s}'_d(l)$ and $\hat{s}''_d(l+1)$ are the known training sequence, (9) is reduced to (12):

$$\Omega(l) = F \left( J \hat{H}(l) \Lambda(l) (J \hat{H}(l))^H + \sigma^2_n JJ^H \right) F^H$$

with approximations (13) and (14) proposed in [7] and [4], respectively:

$$\Delta(l) = \frac{1}{N_d} \left( 1 - E \|s_d(l)\|^2 \right) I_{N_d} \approx FAF^H,$$

$$\sigma^2_n = \frac{1}{N_d} \text{tr}(JJ^H) I_{N_d} \approx \sigma^2_n FJJ^H F^H.$$

Similarly, $\Gamma(l) \in \mathbb{C}^{N_d \times N_d}$ is approximated by (16),

$$\Gamma(l) = \text{diag} \left[ (J \hat{H}(l))^H \Sigma(l)^{-1} J \hat{H}(l) \right]$$

and

$$\hat{r}_d(l) \in \mathbb{C}^{N_d}$$

is

$$\hat{r}_d(l) = r_d(l) - \hat{r}_d(l)$$

and

$$J \hat{y}_d(l) - J \hat{y}_d(l),$$

where

$$\hat{y}_d(l) = \hat{H}(l) \hat{s}_d(l) + \hat{H}'(l) \hat{s}'_d(l) + \hat{H}''(l+1) \hat{s}''_d(l+1).$$

We assume the final output of CHATUE1 $z_1(l)$ can be approximated as an equivalent Gaussian channel output [8], [9] having input $s_d(l)$, as

$$z_1(l) = \mu_{z_1}(l)s_d(l) + n_{z_1}(l),$$

where

$$\mu_{z_1}(l) = \frac{1}{N_d} \text{tr} \left\{ E[z_1(l)s_d(l)] \right\}$$

and

$$n_{z_1}(l) \sim \mathcal{CN}(0, \sigma^2_{z_1}(l))$$

with

$$\sigma^2_{z_1}(l) = \mu_{z_1}(l)(1 - \mu_{z_1}(l)).$$

We finally convert the equalizer output $z_1(l)$ into its corresponding extrinsic LLR, as

$$\lambda_{EQU}(l) = \frac{4 \Re(z_1(l))}{1 - \mu_{z_1}(l)},$$

where $\Re(z_1(l))$ denotes the real part of the complex vector $z_1(l)$.
D. Noise Enhancement with CHATUE1

By utilizing the J-matrix, CHATUE1 has the potential to improve the spectral- and/or energy-efficiencies while keeping the computational complexity order equivalent to that of TEQ-CP. However, CHATUE1 inevitably incurs a noise enhancement problem, as shown in this subsection.

After enough iterations, we can assume $E[|\hat{s}(l)|^2] \rightarrow 1$ at a certain SNR.\(^3\) The mean (21) converges to

$$\mu_{z1} \rightarrow \frac{N_d}{N_d + (N_d + L)\sigma_n^2},$$

as described in Appendix. The variance of the equivalent Gaussian channel output (22) also converges into

$$\sigma_{z1}^2 \rightarrow \frac{N_d(N_d + L)\sigma_n^2}{(N_d + (N_d + L)\sigma_n^2)^2}.$$  \hspace{1cm} (24)

According to [2], the mean $\mu_{z,CP}$ and the variance $\sigma_{z,CP}^2$ of the output of TEQ-CP converge into:

$$\mu_{z,CP} \rightarrow 1 + \frac{\sigma_n^2}{\sigma_n^2}.$$ \hspace{1cm} (26)

$$\sigma_{z,CP}^2 \rightarrow \frac{\sigma_n^2}{(1 + \sigma_n^2)^2}.$$ \hspace{1cm} (27)

respectively, when $E[|\hat{s}_d(l)|^2] \rightarrow 1$.

The asymptotic SNR, $\text{SNR}_{z1}$, of the equalizer output with CHATUE1 is reduced to

$$\text{SNR}_{z1} = \frac{\mu_{z1}^2}{\sigma_{z1}^2} \rightarrow \frac{N_d}{(N_d + L)\sigma_n^2}. $$ \hspace{1cm} (28)

Similarly, the asymptotic SNR, $\text{SNR}_{z,CP}$, of the equalizer output with TEQ-CP is reduced to

$$\text{SNR}_{z,CP} = \frac{\mu_{z,CP}^2}{\sigma_{z,CP}^2} \rightarrow \frac{1}{\sigma_n^2}. $$ \hspace{1cm} (29)

The SNR ratio at the equalizer output of CHATUE1 to that of TEQ-CP is, hence,

$$\frac{1}{2} \leq \frac{\text{SNR}_{z1}}{\text{SNR}_{z,CP}} = \frac{N_d}{N_d + L} \leq 1.$$ \hspace{1cm} (30)

The inequality (30) is because $N_d \geq L \geq 0$. The final output (6) of CHATUE1, thereby, suffers from the noise enhancement of up to 3 dB over TEQ-CP as the IBI length $L$ increases.

E. CHATUE version 2 (CHATUE2)

A motivation of utilizing the $J$-matrix in CHATUE1 is to reduce the computational complexity by restoring the circulant structure of the channel matrix. Although $\mathbf{H} \in \mathbb{C}^{(N_d + L) \times N_d}$ is a Toeplitz matrix, $\mathbf{JH} \in \mathbb{C}^{N_d \times N_d}$ becomes a circulant matrix. Thereby, it is possible to reduce the complexity by exploiting frequency domain processing, since $\mathbf{FJHF}^H$ is a diagonal matrix. On the other hand, CHATUE1 incurs the noise enhancement problem due to the exploitation of the $J$-matrix, as detailed in Section III-D. To cope with the noise enhancement problem, we propose CHATUE2 by introducing a new circulant property restoration method, as follows.

$$\mathbf{r}_d(l) \approx \mathbf{r}_d(l) \triangleq \mathbf{J}_L(1 - \beta)\mathbf{y}_d(l) + \mathbf{G}_L(\beta)\hat{y}_d(l)$$ \hspace{1cm} (31)

$$= \begin{bmatrix} y_d(N_d + 1; l) \\ \vdots \\ y_d(L; l) \end{bmatrix} + \hat{y}_d(N_d + 1; l, \beta),$$ \hspace{1cm} (32)

where $N_d \times (N_d + L)$ matrices $\mathbf{J}_L$ and $\mathbf{G}_L$ are respectively defined as

$$\mathbf{J}_L(1 - \beta) = \begin{bmatrix} \mathbf{O}_{(N_d - L) \times L} \\ (1 - \beta)\mathbf{I}_L \end{bmatrix}_{N_d},$$ \hspace{1cm} (33)

$$\mathbf{G}_L(\beta) = \begin{bmatrix} \mathbf{O}_{N_d} \\ \mathbf{O}_{(N_d - L) \times L} \end{bmatrix}_{\beta L}.$$ \hspace{1cm} (34)

Note that the original $\mathbf{J}$-matrix (8) is identical to $\mathbf{J}_L(1)$. The composite replica $\hat{y}(k; l, \beta)$ is defined as

$$\hat{y}(k; l, \beta) = (1 - \beta)y_d(k; l) + \beta\hat{y}_d(k; l).$$ \hspace{1cm} (35)

We define the factor $\beta$ such that the mean-square-error (MSE) between $\hat{y}(k; l, \beta)$ and $\mathbf{c}_d(l) = \mathbf{H}_d(l)s_d(l) + \mathbf{H}_d'(l)s_d'(l) + \mathbf{H}_d''(l)s_d''(l)$ is minimized, which can be formulated as

$$\beta = \arg \min_{\beta} E \left[ ||\mathbf{c}_d(l) - \hat{y}_d(l, \beta)||^2 \right],$$ \hspace{1cm} (36)

where $\hat{y}_d(l, \beta)$ is the vector version of (35), defined as

$$\hat{y}_d(l, \beta) = (1 - \beta)y_d(l) + \beta\hat{y}_d(l).$$

By taking into account that $E \left[ ||\mathbf{c}_d(l) - \hat{y}_d(l, \beta)||^2 \right] \geq 0$, the problem (36) can be reduced by solving

$$\frac{\partial}{\partial \beta} E \left[ ||\mathbf{c}_d(l) - \hat{y}_d(l, \beta)||^2 \right] = 0.$$ \hspace{1cm} (37)

Since $\mathbf{c}_d(l) = y_d(l) - \mathbf{n}_d$, the solution to (36) is, therefore,

$$\beta = \frac{\sigma_n^2}{E[||y_d(l) - \hat{y}_d(l)||^2]}.$$ \hspace{1cm} (38)

Accordingly, we rewrite (12) as

$$\Omega(l) = \mathbf{F} \left\{ \mathbf{J}_L(1)\mathbf{H}_l\mathbf{A}_l(l)(\mathbf{J}_L(1)\mathbf{H}_l(l))^H \\
+ \sigma_n^2\mathbf{J}_L(1 - \beta)\mathbf{J}_L(1 - \beta)^H \right\} \mathbf{F}^H.$$ \hspace{1cm} (39)

$$\approx \mathbf{F}(l)\mathbf{A}_l(l)(\mathbf{F}(l))^H + \frac{\sigma_n^2 N_d + (1 - \beta)L}{N_d} \mathbf{I}_{N_d}. $$ \hspace{1cm} (40)

The proposed CHATUE2 using (31) and (40) is expected to have the following advantageous points: At the first iteration, (31) is totally equivalent to the original $\mathbf{r}_d(l) = \mathbf{J}_L(1)\mathbf{y}_d(l)$ and CHATUE2 works exactly in the same way as in CHATUE1. After enough iterations are performed, it is expected to satisfy both $\beta \rightarrow 1$ and $E \left[ ||\mathbf{h}(l) - \hat{\mathbf{h}}(l)||^2 \right] < \epsilon + \text{MSE}_{\text{CRB}}$ with an arbitrary small positive value $\epsilon$. The lower bound of the estimation accuracy $\text{MSE}_{\text{CRB}}$ (64) is described in Section V-B. The channel matrix in $\mathbf{r}_d$ approaches
where the burst index $l$ is omitted for the sake of simplicity. The convergence $n \to 1$ contributes to reducing the noise variance (22), through (16), (21) and (40). The mean $\mu_{z2}$ and the variance $\sigma^2_{z2}$ of the equalizer output with $\text{CHATUE2}$, respectively, converge into:

$$
\mu_{z2} \to \frac{N_d}{N_d + (1 - \beta)L} \sigma_n^2, \\
\sigma^2_{z2} \to \frac{N_d(N_d + (1 - \beta)L)\sigma_n^2}{\left(N_d + (1 - \beta)L\right)\sigma_n^2},
$$

when $\text{E}$$[\|\hat{\mathbf{s}}(l)\|^2] \to 1$. Thereby, $\text{CHATUE2}$ improves the signal to noise power ratio $\text{SNR}_{z2}$ at the final equalizer output and it approaches that with $\text{TEQ-CP}$ when $\beta \to 1$, as

$$
\text{SNR}_{z2} \approx \frac{\mu_{z2}^2}{\sigma^2_{z2}} \to \frac{N_d}{N_d + (1 - \beta)L} \frac{\sigma_n^2}{\sigma_n^2} = \frac{1}{\sigma_n^2} = \text{SNR}_{z,\text{CP}}.
$$

IV. Channel Estimation

This section, first of all, reviews the conventional channel estimation techniques [10] briefly. The conventional channel estimation techniques assume $N_t \geq 2W$ when $N_{G1} = N_{G2} = 0$, such that the input signal to the estimator does not suffer from IBI, as we can observe from input data range-$A$ for channel estimation as shown in Fig. 2. Obviously, the longer the training sequence we employ, the lower the spectrum efficiency we have. To improve the spectrum efficiency, we propose a new channel estimation technique, chained turbo estimation (CHATES), which requires $N_t = W$ only. It should be emphasized that the proposed technique is expected to improve the spectrum efficiency without sacrificing the estimation accuracy, and can be applied to $\text{CHATUE1}$, $\text{CHATUE2}$ and $\text{TEQ-CP}$. Note that from the result of the technique described in the previous sections, the chained structure may well be eliminated, and hence the latency problem vanishes. However, the chained structure plays key role when $\text{CHATUE}$ algorithms perform sequence and channel estimation jointly.

A. Review of Channel Estimation Techniques

1) Single Burst ML Channel Estimation: With (46) and (47), single burst maximum likelihood (ML) channel estimation (SB_ML) [10] for the $i$-th iteration is reduced to:

$$
\hat{h}_{SB}^{[i]}(l) = R_{XX}^{[i]}(l)^{-1} R_{XY}^{[i]}(l),
$$

$$
R_{XX}^{[i]}(l) = X(l)^H X(l) + \gamma^{[i-1]}(l) X_d^{[i-1]}(l)^H X_d^{[i-1]}(l),
$$

$$
R_{XY}^{[i]}(l) = X(l)^H y(l) + \gamma^{[i-1]}(l) X_d^{[i-1]}(l)^H y_d(l),
$$

where the $X(l) \in C^{(N_r-W+1)\times W}$ and $X_d^{[i-1]}(l) \in C^{(N_r-W+1)\times W}$ are Toeplitz matrices for the training sequence and the soft replicas of the data symbols, whose first column vectors are $x(l)_{W:N_t}$ and $X_{d,EST}^{[i-1]}(l)_{W:N_t}$, respectively. The soft replica symbol vector $x_{d,EST}^{[i-1]}$ is generated with the LLR of the transmitted data information fed back from the decoder.

We define $x_{d,EST}^{[i-1]}(l) = 0$ for the first iteration ($i = 1$).

$$
\gamma^{[i-1]}(l) = \sigma_n^2 / (\sigma_n^2 + \Delta \sigma_d^{[i-1]}(l)^2),
$$

where $\Delta \sigma_d^{[i-1]}(l)^2 = 1 - \text{E}[\|\hat{x}_d(l)\|^2]$. The soft replica vector $y(l)$ and $y_d(l)$ are respectively defined as $y(l) = [y_0(l), ..., y(N_t-1)]^T \in C^{N_r-W+1}$ and $y_d(l) = [y_{D0} + W, ..., y_{D0} + N_t] \in C^{N_r-W+1}$, where $D_0 = N_t + N_{CP} + N_{G1}$ is the timing offset, in symbols, of the data section.

2) Multi Burst ML Channel Estimation: It is well-known that multi-burst ML channel estimation (MB_ML) [10] improves the estimation accuracy. MB_ML uses a subspace projection technique, and can be approximated by (48) under the assumption that the transmitted symbols are random and long enough.

$$
\hat{h}_{MB}^{[i]}(l) \approx \hat{P}^{[i]}(l) \cdot \hat{h}_{SB}^{[i]}(l).
$$

The projection matrix is given by $\hat{P}^{[i]}(l) = U_p^{[i]}(l) \cdot U_p^{[i]}(l)^H$. The matrix $U_p^{[i]}(l)$ is composed of the first $r$ dominant eigen vectors in $C_h(l)$, referred to as the sample covariance matrix of $\hat{h}$ for the last $L_{MB}$ bursts, as detailed in [10]:

$$
C_h^{[i]}(l) = \frac{1}{L_{MB}} \left\{ \hat{h}_{SB}^{[i]}(l) \hat{h}_{SB}^{[i]}(l)^H \right\} + \sum_{j=-L_{MB}+1}^{(-1)} \hat{h}_{SB}^{[j]}(l) \cdot \hat{h}_{SB}^{[j]}(l)^H,
$$

where $N_t$ denotes the maximum number of iterations. The number $r$ of dominant eigen vectors may be determined using the minimum description length (MDL) [11] for the singular values of $C_h(l)$.

$$
UDV^H = \text{svd} \left( C_h(l) \right),
$$

$$
\hat{r} = \arg \min_{r \leq W} \text{MDL} \{ \text{diag}(D) \}_{1:r}.
$$

B. Chained Turbo Estimation (CHATES)

Turbo channel estimation can estimate the CIR accurately even though the TR length is short, since it extends the reference signal by utilizing the LLR of the transmitted data, fed back from the decoder. Obviously, a shorter TR is preferable from the viewpoint of the spectrum efficiency. In practice, the TR length is designed such that $N_t \geq W$ to estimate a length $W$ CIR. With $N_t = W$, however, the estimation accuracy is degraded because the input signal to the estimator suffers
from IBI, as we can observe from the input data range- 
A as shown in Fig. 2 when $N_{c1} = N_{c2} = 0$. It should be noted that we have to use the input data range- 
A because we can not estimate the length $W$ CIR with the input data range-
A as shown in Fig. 2, since $M^{\text{EST}} < W$ when $N_t = W$, 
where $M^{\text{EST}}$ denotes the length of the input signal to the 
channel estimation with TR. To cope with this problem, we propose a new turbo channel estimation technique, CHATES, 
which performs IBI cancellation for channel estimation. The 
proposed technique is based on the concept of CHATUE 
and improves the spectrum efficiency without sacrificing the 
estimation accuracy. It should be noted, however, the CHATES 
can be applied to the transmission format with a CP as well.

The received training sequence in the current burst $y_t(l) \in \mathbb{C}^{N_t+W-1}$ can be described in the same way as that in (1), 
as:

$$y_t(l) = \mathbf{H}_t(l)s_t(l) + \mathbf{H}'_t(l-1)s'_t(l-1) + \mathbf{H}''_t(l)s''_t(l-1) + \mathbf{n}_t,$$  \hspace{1cm} (52)

where

$$s_t(l) = x_t(l) \quad \in \mathbb{C}^{N_t},$$  \hspace{1cm} (53)

$$s'_t(l-1) = x_{d}(l-1)_{(N_d-N_t+1):N_d} \in \mathbb{C}^{N_t},$$  \hspace{1cm} (54)

and

$$s''_t(l) = x_{d}(l)_{1:N_t} \in \mathbb{C}^{N_t},$$  \hspace{1cm} (55)

if $N_{\text{CP}} = 0$. Otherwise, $s''_t(l)$ indicates the data portion 
including the CP in the current burst,\footnote{In the case of $N_d = W$, $s''_t(l)$ is defined as the CP section only: $s''_t(l) = x_{d}(l)_{(N_d-W+1):N_d}$.} as

$$s''_t(l) = \begin{bmatrix} x_{d}(l)_{(N_d-W+1):N_d} \\ x_{d}(l)_{1:(N_d-W)} \end{bmatrix} \in \mathbb{C}^{N_t}. \hspace{1cm} (56)$$

The matrices $\mathbf{H}_t(l)$, $\mathbf{H}'_t(l-1)$ and $\mathbf{H}''_t(l) \in \mathbb{C}^{(N_t+W-1) \times N_t}$ 
are defined in the same way as (2), (3) and (4), utilizing CIR 

We define an IBI cancelled version of the received training 
sequence $\tilde{y}_t^{[i]} \in \mathbb{C}^{N_t+W-1}$ for the current burst $l$ at $i$-th 
iteration as follows:

$$\tilde{y}_t^{[i]}(l) = y_t(l) - \left\{ \mathbf{H}'_t^{[i-1]}(l-1) s'_t^{[i-1]}(l-1) + \mathbf{H}''_t^{[i-1]}(l) s''_t^{[i-1]}(l) \right\},$$  \hspace{1cm} (57)

where $\mathbf{H}'_t^{[i-1]}(l-1)$, $s'_t^{[i-1]}(l-1)$, $\mathbf{H}''_t^{[i-1]}(l)$ and $s''_t^{[i-1]}(l)$ are obtained as the result of the $(i-1)$-th iteration. We 

We initialize $\mathbf{H}'_t^{[i]}(l-1) = \mathbf{O}$ and $s''_t^{[i]}(l-1) = \mathbf{O}$ for the first iteration ($i = 1$). For the burst 
lased at the head of the frame, we may exploit the result 

The parameters used in the following simulations are de-
tailed in Table I. The Burst Format 1 is used for both 
CHATUE1 and CHATUE2, whereas Burst Format 2 is used for 
TEQ-CP. In the CHATUE algorithms, a data frame encoded 
by a convolutional code $(g_1, g_2) = (7, 5)^8$ with code rate 
$R_c = 1/2$ was divided into $N_B = 10$ bursts. The information 
bits in TEQ-CP, the length of which is the same as the one in 
CHATUE, is encoded with a code with rate $R_c = 2/3$ derived 
from a half rate mother convolutional code $(g_1, g_2) = (7, 5)^8$ 
by puncturing with a puncturing matrix of

$$P_x = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \hspace{1cm} (63)$$

It should be noted the spectrum efficiency is $\eta = 0.4$ in both 
Burst Format 1 for CHATUE algorithms and Burst Format 2 
for TEQ-CP. Thereby, the following comparisons are fair.
### TABLE I

<table>
<thead>
<tr>
<th>Format No.</th>
<th>$N_t$</th>
<th>$N_{G1}$</th>
<th>$N_{CP}$</th>
<th>$N_d$</th>
<th>$N_{G2}$</th>
<th>$R_c$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>0</td>
<td>0</td>
<td>256</td>
<td>0</td>
<td>1/2</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>64</td>
<td>192</td>
<td>0</td>
<td>23/4</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>64</td>
<td>0</td>
<td>256</td>
<td>64</td>
<td>1/2</td>
<td>0.29</td>
</tr>
</tbody>
</table>

A. EXIT Analysis

This subsection shows the results of convergence property analysis of CHATUE2 using EXIT charts. Burst Format 1 described in Table I was used for both CHATUE1 and CHATUE2, whereas Burst Format 2 was used for TEQ-CP.

Fig. 3 shows EXIT curves of CHATUE1 and CHATUE2 as well as TEQ-CP. The equalizer’s EXIT curves were obtained, in all the system setups tested, for a 64-path frequency selective Rayleigh fading channel realization with average SNR = 2.4 dB. Ideal channel estimation is assumed. The mutual information (MI) $I_{EQU}^\lambda$ between the LLR $\lambda_{EQU}^c$ and the coded bits input to the symbol mapper $c_M$ is defined by

$$I_{EQU}^\lambda = I(\lambda_{EQU}^c; c_M) = \frac{1}{2} \sum_{m=\pm 1} \int_{-\infty}^{+\infty} \Pr(\lambda_{EQU}^c|m) \log_2 \frac{\Pr(\lambda_{EQU}^c|m)}{\Pr(\lambda_{EQU}^c)} d\lambda_{EQU}^c,$$

where $\Pr(\lambda_{EQU}^c|m)$ is the conditional probability density of $\lambda_{EQU}^c$ given $m = 1 - 2c_M$ [9].

It is found from Fig. 3 that the equalizer’s EXIT curve of CHATUE1 is located below the TEQ-CP’s EXIT curve over entire value range of a priori mutual information $I_{EQU}^a$. This is because of the noise enhancement described in Section III-D. In contrast, CHATUE2 improves $I_{EQU}^\lambda$ and achieves almost the same point as that with TEQ-CP when $I_{EQU}^\lambda = 1$, although its left most point at $I_{EQU}^\lambda = 0$ is almost the same as that of CHATUE1. This observation verifies the asymptotic perfect elimination of the noise enhancement with the CHATUE2 algorithm.

A trajectory of turbo equalization with CHATUE2 is also presented in Fig. 3. The trajectory reaches a point very close to $I_{DEC} = 1$ without intersection in the channel realization used and hence the MI between the a posteriori LLR of decoder $\lambda_{DEC}^b$ and the binary source information approaches 1. This is because of two reasons: 1) CHATUE2 improves the equalizer’s EXIT curve by eliminating the noise enhancement; 2) CHATUE2 algorithms allow us to use a lower rate code by utilizing the time duration allocated for CP. On the other hand, the EXIT curves of CHATUE1 and TEQ-CP has the intersection at (0.98, 0.8) and (0.92, 0.85), respectively. Thereby the trajectories of CHATUE1 and TEQ-CP can rarely approach points very close to $I_{DEC} = 1$ for a SNR of 2.4 dB, although they are not presented in Fig. 3 to avoid too dense a representation. This is because CHATUE1 incurs the noise enhancement at the equalizer output or TEQ-CP can not use a lower rate code with the same spectrum efficiency due to CP-transmission.

B. Performance of CHATES

This subsection presents results of simulations conducted to verify the channel estimation accuracy improvement achieved by the proposed joint IBI cancellation and channel estimation technique, CHATES. The two transmission formats, Burst Format 1 and 3 described in Table I were assumed because of the reasons as follows: Burst Format 1, which can cause IBI in the TR section, is used to verify the proposed IBI cancellation technique; Burst Format 3 has guard intervals on both the sides in time of TR such that the TR section does not suffer from IBI due to the neighboring data sections. The reason for presenting the simulation result with Burst Format 3 is to provide a basis for the performance comparison of the IBI cancellation, although its spectrum efficiency is less than that of Burst Format 1. The parameter $L_{MB}$ of MB_ML was set at 300.

Fig. 4 shows the MSE of the channel estimate with CHATUE2 in a six path fading channel realization based on the pedestrian-B model [12] with a 3 km/h (PB3) mobility assumption. The path positions are at {0, 3, 12, 18, 34.5, 55.5} symbol timings assuming that a transmission bandwidth of 15 MHz. We have $-4 \text{ dB} \leq \text{SNR} \leq 16 \text{ dB}$ which corresponds to $0 \text{ dB} \leq E_b/N_0 \leq 20 \text{ dB}$ with Burst Format 1. It is found that CHATES, with six iterations, asymptotically achieves the equivalent MSE to the analytical accuracy bound of MB_ML with $L_{MB} \rightarrow \infty$ [10], given by

$$\text{MSE}_{\text{CRB}}(\sigma_0) = \frac{r}{M_t \text{est} + M_d \text{est}} \sigma_n^2, \quad (64)$$

where $r$ is the number of dominant paths. $M_t \text{est}$ and $M_d \text{est}$ indicate the length of the input data to MB_ML for the training and the data section, respectively. Note that the bound (64) is equivalent to the CRB as described in [10]. Without the IBI cancellation technique, the estimation accuracy degrades due to IBI even after six iterations are performed.
C. BER performance

This subsection presents the BER performance of CHATUE2, in comparison to CHATUE1 and TEQ-CP. Burst Format 1 described in Table I was used for both CHATUE1 and 2, whereas Burst Format 2 was used for TEQ-CP.

In Fig. 5, the BER performance of turbo equalization for a single path static AWGN channel are presented, even though equalization is not needed in single path channels. This is because the purpose of showing the BER performance with the known channel is to make a baseline comparison of the techniques: CHATUE1, 2 and TEQ-CP. For reference, the BER performance of BCJR decoders with the parameters mentioned above are also presented. The BER with TEQ-CP is the same as that with a) BCJR decoder ($R_c = 2/3$) used in TEQ-CP, as shown in Fig. 5. However, the BER with CHATUE1 is degraded compared to c) BCJR decoder ($R_c = 1/2$) due to the noise enhancement described in Section III-D. The BER with CHATUE1 is identical to that with b) BCJR decoder ($R_c = 1/2$) with the noise enhancement to its input before interleaving. The noise enhancement localized in the $L$ symbols is not uniformly distributed over a frame even after interleaving and hence it degrades the performance of a BCJR decoder more than expected (0.97 dB), as shown in (30). The BER with CHATUE2, on the other hand, achieves exactly the same as that with c) BCJR decoder ($R_c = 1/2$), in the same way as for TEQ-CP. It should be noted that the noise power of input signal to the BCJR decoder b) is intentionally enhanced to reproduce the noise enhancement problem incurred by CHATUE1. The noise power of the input signal to the BCJR decoder b) is increased to $2\sigma^2$ for the first $L$ bits. The BCJR decoder b) decodes the noise enhanced input signal following interleaving. The BCJR decoder b) itself is the same as BCJR decoder c).
The primary objective of this paper has been to provide solutions to the problems inherent in chained turbo equalization techniques, which are: 1) the latency due to the time-concatenation of equalization, and 2) the noise enhancement at the equalizer output. This paper showed that Problem 1) can easily be solved with a practical and reasonable assumption that the training sequence is transmitted in every burst. To cope with Problem 2), this paper proposed chained turbo equalization version 2, CHATUE2. Since CHATUE2 utilizes the composite replica to retrieve the circulant structure of the channel matrix in the received signal, CHATUE2 improves the equalizer output SNR to the same level as that with TEQ-CP.

Furthermore, this paper proposed a new IBI cancellation technique for channel estimation, chained turbo estimation (CHATUE), that improves spectrum efficiency without sacrificing estimation accuracy. CHATUE can be applied to CHATUE1, CHATUE2 and TEQ-CP, although CHATUE inherits the CHATUE concept in the sense that the cancellation of IBI occurring in the TR section utilizes the LLR of transmitted data, fed back from the decoder, not only in the current but also in the past bursts.

The results of computer simulations showed that CHATUE achieves asymptotically equivalent estimation accuracy to the CRB with a short training sequence, the length of which is equal to the length of channel impulse response. Results of BER simulations were also presented in this paper to demonstrate the effectiveness of the proposed technique in realistic scenarios based on measurement data, as well as in model-based frequency selective fading channels. The simulation results showed that CHATUE2 further improves the BER over CHATUE1 and achieves a gain of more than 1 dB over TEQ-CP at BER = 10^{-5}, when the same spectrum efficiency is assumed for the equalization techniques.

**Appendix**

### A. Derivation of the asymptotic mean (24)

Assuming that $E[|\hat{s}(l)|^2] \to 1$ after enough iterations, the noise covariance matrix (12) converges to

$$
\mathbf{\Omega}(l) \to \sigma_n^2 \mathbf{N}_d + \frac{L}{\mathbf{N}_d} \mathbf{I}_{\mathbf{N}_d}.
$$

Hence, the equation (16) converges to

$$
\Gamma(l) \to \frac{\mathbf{N}_d}{(\mathbf{N}_d + L)\sigma_n^2} \mathbf{I}_{\mathbf{N}_d},
$$

under the assumption $E[|\mathbf{h}(l)|^2] = E[|\mathbf{h}(l)|^2] = 1$. The asymptotic mean (24) is reduced by substituting (65), (66) and $\mathbf{S}(l) \to \mathbf{I}_{\mathbf{N}_d}$ into (21).

**Acknowledgment**

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, which are very helpful to improve the quality of this paper.
REFERENCES


Yasuhiro Takano received BS degree from the department of Mathematics, Rikkyo University, Japan in 1999. He received MS degree from the school of Information science, Japan Advanced Institute of Science and Technology (JAIST) in 2010. Now he is a Ph.D. student in Information theory and signal processing laboratory in JAIST. His research interests are channel estimation and equalization.

Khoirul Anwar graduated (cum laude) from the department of Electrical Engineering (Telecommunications), Institut Teknologi Bandung (ITB), Bandung, Indonesia in 2000. He received Master and Doctor Degrees from Graduate School of Information Science, Nara Institute of Science and Technology (NAIST) in 2005 and 2008, respectively. Since then, he has been appointed as an assistant professor in NAIST. He received best student paper award from the IEEE Radio and Wireless Symposium 2006 (RWS06), California-USA, and Best Paper of Indonesian Student Association (ISA 2007), Kyoto, Japan in 2007. Since September 2008, he is with the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST) as an assistant professor. His research interests are network information theory, error control coding, iterative decoding and signal processing for wireless communications. He has authored around 58 scientific publications in these areas. He serves as a reviewer for a number of main journals and conferences in the area of wireless communications and signal processing. Dr. Anwar is a member of IEEE, information theory society, and IEICE Japan.