JAIST Repository
https://dspace.jaist.ac.jp/

Title	Spectrally Efficient Frame Format-Aided Turbo Equalization with Channel Estimation
Author(s)	Takano, Yasuhiro; Anwar, Khoirul; Matsumoto, Tad
Citation	IEEE Transactions on Vehicular Technology, 62(4): 1635-1645
Issue Date	2012-11-30
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/10859
Rights	This is the author's version of the work. 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.
Description	

Japan Advanced Institute of Science and Technology

Spectrally Efficient Frame Format–Aided Turbo Equalization with Channel Estimation

Yasuhiro Takano, *Student Member, IEEE,* Khoirul Anwar, *Member, IEEE,* and Tad Matsumoto, *Fellow, IEEE*

***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

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Y. Takano, K. Anwar and T. Matsumoto are with the Japan Advanced Institute of Science and Technology (JAIST) 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan (e-mail: {yace.takano; anwar-k; matumoto}@jaist.ac.jp). T. Matsumoto is also with the Centre for Wireless Communications, University of Oulu, FIN-90014, Oulu, Finland.

This research was conducted under the financial support of, in part, the double degree program between JAIST and University of Oulu, and in part by the Japan Society for the Promotion of Science (JSPS) Grant under the Scientific Research (C) No. 22560367.

$N_t \geq 2W$ 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, CHATES. 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 $\mathbf{X}$ and lower-case $\mathbf{x}$ denote a matrix and a vector, respectively. $\mathbf{X}^H$ denotes the transposed conjugation of the matrix $\mathbf{X}$. $\mathtt{diag}(\mathbf{X})$ is an operator that forms a vector from the diagonal elements of its argument matrix $\mathbf{X}$, while $\mathtt{DIAG}(\mathbf{x})$ forms a diagonal matrix from its argument vector $\mathbf{x}$. $\mathtt{svd}(\mathbf{X}) = \mathbf{UDV}^H$ is the singular value decomposition of a matrix $\mathbf{X} \in \mathbb{C}^{M\times N}$, where $\mathbf{U} \in \mathbb{C}^{M\times M}$ and $\mathbf{V} \in \mathbb{C}^{N\times N}$ are unitary matrices and $\mathbf{D} \in \mathbb{C}^{M\times N}$ is a rectangular diagonal matrix. $\mathbf{X}|_{1:r}$ is a submatrix composed of the first r column vectors in a matrix $\mathbf{X}$. Similarly, $\mathbf{x}|_{i:j}$ is a subvector of the original vector $\mathbf{x}$ which extracts the i-th to the j-th elements from the vector $\mathbf{x}$.

II. System Model

The system model assumed in this paper is depicted in Fig. 1. A binary data information sequence $b(i)$, $1 \leq i \leq N_B N_d R_c$, is encoded by a rate R_c convolutional code (CC) with generator polynomials $(g_1, \cdots, g_{1/R_c})$ and is interleaved by an interleaver (Π). 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[1] $x(k_s; l)$ 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_{G1} 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 inter-symbol-interference (ISI) due to fading frequency selectivity, 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 $\mathbf{y}(l) \in \mathbb{C}^{K+L}$, as

$$\begin{aligned} \mathbf{y}(l) &= \mathbf{H}(l)\mathbf{x}(l) + \mathbf{H}'(l-1)\mathbf{x}'(l-1) \\ &\quad + \mathbf{H}''(l+1)\mathbf{x}''(l+1) + \mathbf{n}, \end{aligned} \tag{1}$$

where $\mathbf{x}(l)$, $\mathbf{x}'(l-1)$ and $\mathbf{x}''(l+1)$ 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. $\mathbf{H}(l) \in \mathbb{C}^{(K+L)\times K}$ is a Toeplitz matrix representing the convolution of the transmitted data symbols with the CIR in the current block, with

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

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

$$\mathbf{H}'(l-1) = \begin{bmatrix} \mathbf{O}_{L\times(K-L)} & \mathbf{H}'_{\nabla}(l-1) \\ \mathbf{O}_{K\times(K-L)} & \mathbf{O}_{K\times L} \end{bmatrix} \tag{3}$$

with

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

and

$$\mathbf{H}''(l+1) = \begin{bmatrix} \mathbf{O}_{K\times L} & \mathbf{O}_{K\times(K-L)} \\ \mathbf{H}''_{\Delta}(l+1) & \mathbf{O}_{L\times(K-L)} \end{bmatrix} \tag{4}$$

with

$$\mathbf{H}''_{\Delta}(l) = \begin{bmatrix} h(1;l) & & & 0 \\ h(2;l) & h(1;l) & & \\ \vdots & \vdots & \ddots & \\ h(L;l) & h(L-1;l) & \cdots & h(1;l) \end{bmatrix}.$$

$\mathbf{n}$ is a complex AWGN vector, the elements of which follow $\mathcal{CN}(0, \sigma_n^2)$, where the variance σ_n^2 is determined according to the SNR.

As depicted in Fig. 1, the receiver performs channel estimation (EST) while also obtaining the extrinsic log-likelihood ratio (LLR) λ^e_{EQU} corresponding to the transmitted sequence $x(k_s; l)$ by means of frequency domain soft-cancellation and minimum mean-square-error (FD/SC-MMSE) turbo equalization [2] (EQU). Using the LLR λ^a_{DEC} after interleaving λ^e_{EQU}, the channel decoder (CC^{-1}) performs decoding by using the Bahl, Cocke, Jelinek and Raviv (BCJR) algorithm [6] and outputs the *a posteriori* LLR (λ^p_{DEC}) corresponding to $c(i)$ which is used to generate the soft replica of

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

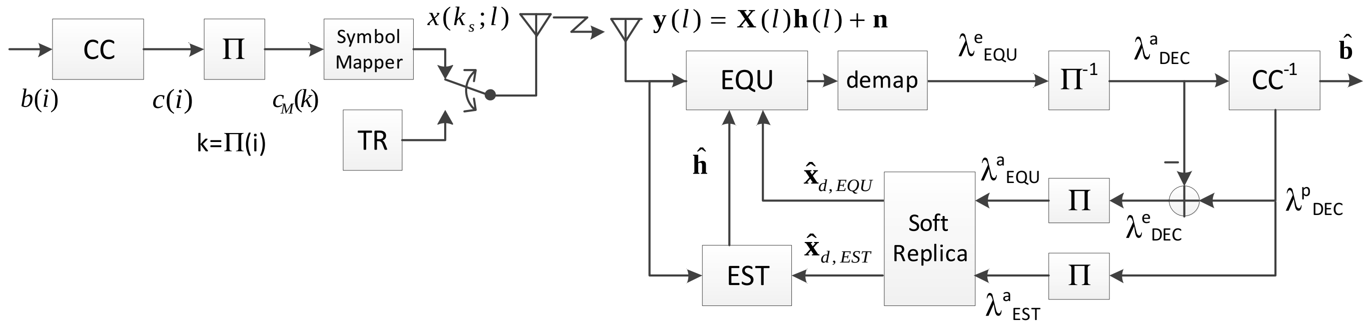

Fig. 1. System model: Structure of transmitter and receiver.

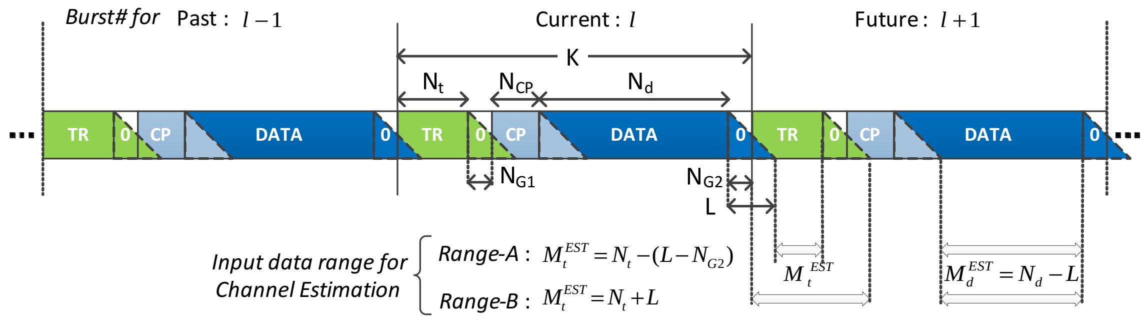

Fig. 2. Structure of burst formats and illustration of input data range for channel estimation.

the transmitted symbols. CC^{-1} outputs the estimates of the transmitted sequence $\hat{\mathbf{b}}$ by making a hard decision on the decoder's *a posteriori* LLR λ^p_{DEC} corresponding to $b(i)$ after several iterations. EST and EQU utilize the soft replica of the transmitted symbols $\hat{\mathbf{x}}_{d,EST}$ and $\hat{\mathbf{x}}_{d,EQU}$, respectively. $\hat{\mathbf{x}}_{d,EST}$ is generated from the *a priori* LLR λ^a_{EST} for channel estimation after interleaving the *a posteriori* LLR λ^p_{DEC}. On the other hand, $\hat{\mathbf{x}}_{d,EQU}$ is generated from the equalizer's *a priori* LLR λ^a_{EQU} which is the interleaved version of the extrinsic LLR $\lambda^e_{DEC} = \lambda^p_{DEC} - \lambda^a_{DEC}$, 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 $N_t = N_{CP} = N_{G1} = N_{G2} = 0$. 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 $N_t > L$ and $N_{CP} = N_{G1} = N_{G2} = 0$, IBIs 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.

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 $\mathbf{y}_d(l) \in \mathbb{C}^{N_d+L}$ for the transmitted burst in the current burst timing l is described as

$$\begin{aligned}\mathbf{y}_d(l) &= \mathbf{H}_d(l)\mathbf{s}_d(l) + \mathbf{H}'_d(l)\mathbf{s}'_d(l) \\ &\quad + \mathbf{H}''_d(l+1)\mathbf{s}''_d(l+1) + \mathbf{n}_d, \end{aligned} \quad (5)$$

where the signal vectors, each of which the size is $N_d \times 1$, are defined as $\mathbf{s}_d(l) = \mathbf{x}_d(l)$, $\mathbf{s}'_d(l) = [\mathbf{0}_{1\times(N_d-W)} \ \ \mathbf{x}_t(l)^T]^T$ and $\mathbf{s}''_d(l+1) = [\mathbf{x}_t(l+1)^T \ \ \mathbf{0}_{1\times(N_d-W)}]^T$. The $(N_d+L)\times N_d$ Toeplitz matrices $\mathbf{H}_d(l)$, $\mathbf{H}'_d(l)$ and $\mathbf{H}''_d(l+1)$ are defined in the same way as (2), (3) and (4), utilizing CIR vectors $\mathbf{h}(l)$, $\mathbf{h}(l)$ and $\mathbf{h}(l+1)$, respectively.[2] The $N_d \times 1$ noise vector $\mathbf{n}_d$ follows $\mathcal{CN}(0, \sigma_n^2)$.

C. Review of CHATUE version 1 (CHATUE1)

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

$$\begin{aligned}\mathbf{z}_1(l) &= \left(\mathbf{I}_{N_d} + \mathbf{\Gamma}(l)\hat{\mathbf{S}}(l)\right)^{-1} \\ &\quad \cdot \left[\mathbf{\Gamma}(l)\hat{\mathbf{s}}_d(l) + \mathbf{F}^H\hat{\mathbf{\Phi}}(l)^H\mathbf{\Omega}(l)^{-1}\mathbf{F}\tilde{\mathbf{r}}_d(l)\right], \end{aligned} \quad (6)$$

where $\hat{\mathbf{S}}(l) = \mathtt{DIAG}[||\hat{\mathbf{s}}_d(l)||^2]$ and

$$\hat{\mathbf{\Phi}}(l) = \mathbf{FJ}\hat{\mathbf{H}}(l)\mathbf{F}^H \quad (7)$$

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

$$\mathbf{J} = \begin{pmatrix} \mathbf{O}_{(N_d-L)\times L} & \\ \mathbf{I}_L & \mathbf{I}_{N_d} \end{pmatrix} \in \mathbb{R}^{N_d\times(N_d+L)}. \quad (8)$$

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

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

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

$$\mathbf{\Omega}(l) = \mathbf{F\Sigma}(l)\mathbf{F}^H, \quad (9)$$

where

$$\begin{aligned}\mathbf{\Sigma}(l) &= \mathbf{J}\hat{\mathbf{H}}(l)\mathbf{\Lambda}(l)(\mathbf{J}\hat{\mathbf{H}}(l))^H \\ &\quad + \mathbf{J}\hat{\mathbf{H}}'(l)\mathbf{\Lambda}'(l)(\mathbf{J}\hat{\mathbf{H}}'(l))^H \\ &\quad + \mathbf{J}\hat{\mathbf{H}}''(l+1)\mathbf{\Lambda}''(l+1)(\mathbf{J}\hat{\mathbf{H}}''(l+1))^H \\ &\quad + \sigma_n^2\mathbf{JJ}^H \end{aligned} \quad (10)$$

[2]As we mentioned above, IBIs for a current data are caused by TRs in the current burst index l or future $l+1$ when $N_{CP} = N_{G1} = N_{G2} = 0$ in Fig. 2. Hence, $\mathbf{H}'_d(l)$ is constructed with $\mathbf{h}(l)$.

with

$$\begin{aligned}\mathbf{\Lambda}(l) &= \mathrm{E}\left[\{\hat{\mathbf{s}}_d(l) - \mathbf{s}_d(l)\}\{\hat{\mathbf{s}}_d(l) - \mathbf{s}_d(l)\}^H\right], \\ \mathbf{\Lambda}'(l) &= \mathrm{E}\left[\{\hat{\mathbf{s}}'_d(l) - \mathbf{s}'_d(l)\}\{\hat{\mathbf{s}}'_d(l) - \mathbf{s}'_d(l)\}^H\right]\end{aligned}$$

and

$$\begin{aligned}\mathbf{\Lambda}''(l+1) &= \mathrm{E}\left[\{\hat{\mathbf{s}}''_d(l+1) - \mathbf{s}''_d(l+1)\}\right. \\ &\quad \left.\cdot\{\hat{\mathbf{s}}''_d(l+1) - \mathbf{s}''_d(l+1)\}^H\right].\end{aligned}$$

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

$$\begin{aligned}\mathbf{\Omega}(l) &= \mathbf{F}\left(\mathbf{J}\hat{\mathbf{H}}(l)\mathbf{\Lambda}(l)(\mathbf{J}\hat{\mathbf{H}}(l))^H + \sigma_n^2\mathbf{JJ}^H\right)\mathbf{F}^H \quad (11) \\ &\approx \hat{\mathbf{\Phi}}(l)\mathbf{\Delta}(l)\hat{\mathbf{\Phi}}(l)^H + \sigma_n^2\frac{N_d+L}{N_d}\mathbf{I}_{N_d}, \quad (12)\end{aligned}$$

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

$$\mathbf{\Delta}(l) = \frac{1}{N_d}\left(1 - \mathrm{E}\left[||\hat{\mathbf{s}}_d(l)||^2\right]\right)\mathbf{I}_{N_d} \approx \mathbf{F\Lambda F}^H, \quad (13)$$

$$\sigma_n^2\frac{N_d+L}{N_d}\mathbf{I}_{N_d} = \sigma_n^2\frac{\mathrm{tr}(\mathbf{JJ}^H)}{N_d}\mathbf{I}_{N_d} \approx \sigma_n^2\mathbf{FJJ}^H\mathbf{F}^H. \quad (14)$$

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

$$\begin{aligned}\mathbf{\Gamma}(l) &= \mathtt{diag}\left[(\mathbf{J}\hat{\mathbf{H}}(l))^H\mathbf{\Sigma}(l)^{-1}\mathbf{J}\hat{\mathbf{H}}(l)\right] \quad (15) \\ &\approx \frac{1}{N_d}\mathrm{tr}\left[\hat{\mathbf{\Phi}}(l)^H\mathbf{\Omega}(l)^{-1}\hat{\mathbf{\Phi}}(l)\right]\mathbf{I}_{N_d}. \quad (16)\end{aligned}$$

The residual $\tilde{\mathbf{r}}_d(l) \in \mathbb{C}^{N_d}$ is

$$\begin{aligned}\tilde{\mathbf{r}}_d(l) &= \mathbf{r}_d(l) - \hat{\mathbf{r}}_d(l) \quad (17) \\ &= \mathbf{Jy}_d(l) - \mathbf{J}\hat{\mathbf{y}}_d(l), \quad (18)\end{aligned}$$

where

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

We assume the final output of *CHATUE1* $\mathbf{z}_1(l)$ can be approximated as an equivalent Gaussian channel output [8], [9] having input $\mathbf{s}_d(l)$, as

$$\mathbf{z}_1(l) = \mu_{z1}(l)\mathbf{s}_d(l) + \mathbf{n}_{z1}(l), \quad (19)$$

where

$$\begin{aligned}\mu_{z1}(l) &= \frac{1}{N_d}\mathrm{tr}\left\{\mathrm{E}[\mathbf{z}_1(l)\mathbf{s}_d^H(l)]\right\} \quad (20) \\ &= \frac{1}{N_d}\mathrm{tr}\left\{(\mathbf{I}_{N_d} + \mathbf{\Gamma}(l)\hat{\mathbf{S}}(l))^{-1}\mathbf{\Gamma}(l)\right\}\mathrm{E}[||\mathbf{s}_d(l)||^2] \quad (21)\end{aligned}$$

and $\mathbf{n}_{z1}(l) \sim \mathcal{CN}(0, \sigma_{z1}^2(l))$ with

$$\sigma_{z1}^2(l) = \mu_{z1}(l)(1 - \mu_{z1}(l)). \quad (22)$$

We finally convert the equalizer output $\mathbf{z}_1(l)$ into its corresponding extrinsic LLR, as

$$\lambda_{EQU}^e(l) = \frac{4\Re(\mathbf{z}_1(l))}{1 - \mu_{z1}(l)}, \quad (23)$$

where $\Re(\mathbf{z}_1(l))$ denotes the real part of the complex vector $\mathbf{z}_1(l)$.

D. Noise Enhancement with CHATUE1

By utilizing the $\mathbf{J}$-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 $\mathrm{E}[||\hat{\mathbf{s}}(l)||^2] \to 1$ at a certain SNR.[3] The mean (21) converges to

$$\mu_{z1} \quad \to \quad \frac{N_d}{N_d + (N_d+L)\sigma_n^2}, \tag{24}$$

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

$$\sigma_{z1}^2 \quad \to \quad \frac{N_d(N_d+L)\sigma_n^2}{\{N_d + (N_d+L)\sigma_n^2\}^2}. \tag{25}$$

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

$$\mu_{z,\mathrm{CP}} \quad \to \quad \frac{1}{1+\sigma_n^2}, \tag{26}$$

$$\sigma_{z,\mathrm{CP}}^2 \quad \to \quad \frac{\sigma_n^2}{\{1+\sigma_n^2\}^2}, \tag{27}$$

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

The asymptotic SNR, SNR_{z1}, of the equalizer output with *CHATUE1* is reduced to

$$\mathrm{SNR}_{z1} = \frac{\mu_{z1}^2}{\sigma_{z1}^2} \quad \to \quad \frac{N_d}{(N_d+L)\sigma_n^2}. \tag{28}$$

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

$$\mathrm{SNR}_{z,\mathrm{CP}} = \frac{\mu_{z,\mathrm{CP}}^2}{\sigma_{z,\mathrm{CP}}^2} \quad \to \quad \frac{1}{\sigma_n^2}. \tag{29}$$

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

$$\frac{1}{2} \quad \leq \quad \frac{\mathrm{SNR}_{z1}}{\mathrm{SNR}_{z,\mathrm{CP}}} \quad = \quad \frac{N_d}{N_d+L} \quad \leq \quad 1. \tag{30}$$

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 $\mathbf{J}$-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 Toepliz 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 $\mathbf{J}$-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.

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

$$= \begin{bmatrix} y_d(L+1;l) & \\ \vdots & \\ y_d(N_d;l) & \\ y_d(1;l) & + \bar{y}_d(N_d+1;l,\beta) \\ \vdots & \\ y_d(L;l) & + \bar{y}_d(N_d+L;l,\beta) \end{bmatrix}, \tag{32}$$

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

$$\mathbf{J}_L(1-\beta) \quad = \quad \begin{pmatrix} \mathbf{O}_{(N_d-L)\times L} & \mathbf{I}_{N_d} \\ (1-\beta)\mathbf{I}_L & \end{pmatrix}, \tag{33}$$

$$\mathbf{G}_L(\beta) \quad = \quad \begin{pmatrix} \mathbf{O}_{N_d} & \mathbf{O}_{(N_d-L)\times L} \\ & \beta\mathbf{I}_L \end{pmatrix}. \tag{34}$$

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

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

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

$$\beta \quad = \quad \arg\min_{\beta} \mathrm{E}\left[||\mathbf{c}_d(l) - \bar{\mathbf{y}}_d(l,\beta)||^2\right], \tag{36}$$

where $\bar{\mathbf{y}}_d(l,\beta)$ is the vector version of (35), defined as $\bar{\mathbf{y}}_d(l,\beta) = (1-\beta)\mathbf{y}_d(l) + \beta\hat{\mathbf{y}}_d(l)$. By taking into account that $\mathrm{E}\left[||\mathbf{c}_d(l) - \bar{\mathbf{y}}_d(l,\beta)||^2\right] \geq 0$, the problem (36) can be reduced by solving

$$\frac{\partial}{\partial\beta}\mathrm{E}\left[||\mathbf{c}_d(l) - \bar{\mathbf{y}}_d(l,\beta)||^2\right] \quad = \quad 0. \tag{37}$$

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

$$\beta \quad = \quad \frac{\sigma_n^2}{\mathrm{E}\left[||\mathbf{y}_d(l) - \hat{\mathbf{y}}_d(l)||^2\right]}. \tag{38}$$

Accordingly, we rewrite (12) as,

$$\begin{aligned} \mathbf{\Omega}(l) \quad &= \quad \mathbf{F}\Big\{\mathbf{J}_L(1)\hat{\mathbf{H}}(l)\mathbf{\Lambda}(l)(\mathbf{J}_L(1)\hat{\mathbf{H}}(l))^H \\ &\qquad + \sigma_n^2\mathbf{J}_L(1-\beta)\mathbf{J}_L(1-\beta)^H\Big\}\mathbf{F}^H & (39) \\ &\approx \quad \hat{\mathbf{\Phi}}(l)\mathbf{\Delta}(l)\hat{\mathbf{\Phi}}(l)^H + \sigma_n^2\frac{N_d+(1-\beta)L}{N_d}\mathbf{I}_{N_d}. & (40) \end{aligned}$$

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 $\beta \to 1$ and $\mathrm{E}\left[||\mathbf{h}(l) - \hat{\mathbf{h}}(l)||^2\right] < \epsilon + \mathrm{MSE}_{\mathrm{CRB}}$ with an arbitrary small positive value ϵ. The lower bound of the estimation accuracy $\mathrm{MSE}_{\mathrm{CRB}}$ (64) is described in Section V-B. The channel matrix in $\bar{\mathbf{r}}_d$ approaches

[3]The required SNR falls into the issue of matching between the equalizer and decoder's EXIT curves. However, it is out of the scope of this paper.

a matrix having a circulant structure when the estimate $\hat{\mathbf{h}}$ is accurate:

$$\mathbf{J}_L(1-\beta)\mathbf{H}+\mathbf{G}_L(\beta)\hat{\mathbf{H}} \overset{\beta\to 1}{=} \begin{bmatrix} h(W) & \cdots & h(2) & h(1) & & & \\ & \ddots & \vdots & h(2) & h(1) & & \\ & & h(W) & \vdots & h(2) & \ddots & \\ & & & h(W) & \vdots & \ddots & h(1) \\ \hat{h}(1) & & & & h(W) & & h(2) \\ \vdots & \ddots & & & & \ddots & \vdots \\ \hat{h}(L) & \cdots & \hat{h}(1) & & & & h(W) \end{bmatrix}, \quad (41)$$

where the burst index l is omitted for the sake of simplicity. The convergence $\beta \to 1$ contributes to reducing the noise variance (22), through (16), (21) and (40). The mean μ_{z2} and the variance σ_{z2}^2 of the equalizer output with *CHATUE2*, respectively, converge into:

$$\mu_{z2} \to \frac{N_d}{N_d+(N_d+(1-\beta)L)\sigma_n^2}, \quad (42)$$

$$\sigma_{z2}^2 \to \frac{N_d(N_d+(1-\beta)L)\sigma_n^2}{\{N_d+(N_d+(1-\beta)L)\sigma_n^2\}^2}, \quad (43)$$

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

$$\begin{aligned} \mathrm{SNR}_{z2} &= \frac{\mu_{z2}^2}{\sigma_{z2}^2} && (44) \\ &\to \frac{N_d}{\{N_d+(1-\beta)L\}\,\sigma_n^2} \xrightarrow{\beta\to 1} \frac{1}{\sigma_n^2} = \mathrm{SNR}_{z,\mathrm{CP}}. \end{aligned}$$

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 *CHATUE1*, *CHATUE2* and 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 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:

$$\hat{\mathbf{h}}_{SB}^{[i]}(l) = \mathbf{R}_{\mathbf{XX}}^{[i]}(l)^{-1}\,\mathbf{R}_{\mathbf{XY}}^{[i]}(l), \quad (45)$$

$$\mathbf{R}_{\mathbf{XX}}^{[i]}(l) = \mathbf{X}_t(l)^H\,\mathbf{X}_t(l)+\gamma^{[i-1]}(l)\,\hat{\mathbf{X}}_d^{[i-1]}(l)^H\,\hat{\mathbf{X}}_d^{[i-1]}(l), \quad (46)$$

$$\mathbf{R}_{\mathbf{XY}}^{[i]}(l) = \mathbf{X}_t(l)^H\,\mathbf{y}_t(l)+\gamma^{[i-1]}(l)\,\hat{\mathbf{X}}_d^{[i-1]}(l)^H\,\mathbf{y}_d(l). \quad (47)$$

$\mathbf{X}_t(l) \in \mathbb{C}^{(N_t-W+1)\times W}$ and $\hat{\mathbf{X}}_d^{[i-1]}(l) \in \mathbb{C}^{(N_d-W+1)\times W}$ are Toeplitz matrices for the training sequence and the soft replicas of the data symbols, whose first column vectors are $\mathbf{x}_t(l)|_{W:N_t}$ and $\hat{\mathbf{x}}_{d,EST}^{[i-1]}(l)|_{W:N_d}$, respectively. The soft replica symbol vector $\hat{\mathbf{x}}_{d,EST}^{[i-1]}$ is generated with the LLR of the transmitted data information fed back from the decoder. We define $\hat{\mathbf{x}}_{d,EST}^{[i-1]}(l) = \mathbf{0}$ for the first iteration ($i = 1$). $\gamma^{[i-1]}(l) = \sigma_n^2/(\sigma_n^2+\Delta\sigma_d^{[i-1]}(l)^2)$ with $\Delta\sigma_d^{[i-1]}(l)^2 = 1 - \mathrm{E}[\|\hat{\mathbf{x}}_d^{[i-1]}(l)\|^2]$. $\mathbf{y}_t(l)$ and $\mathbf{y}_d(l)$ are respectively defined as $\mathbf{y}_t(l) = [y(W;l),\cdots,y(N_t;l)]^T \in \mathbb{C}^{N_t-W+1}$ and $\mathbf{y}_d(l) = [y(D_0+W;l),\cdots,y(D_0+N_d;l)]^T \in \mathbb{C}^{N_d-W+1}$, 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:

$$\hat{\mathbf{h}}_{MB}^{[i]}(l) \approx \hat{\mathbf{P}}^{[i]}(l)\cdot\hat{\mathbf{h}}_{SB}^{[i]}(l). \quad (48)$$

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

$$\begin{aligned} \mathbf{C}_{\hat{\mathbf{h}}}^{[i]}(l) &= \frac{1}{L_{MB}}\Bigg\{\hat{\mathbf{h}}_{SB}^{[i]}(l)\,\hat{\mathbf{h}}_{SB}^{[i]}(l)^H \\ &\quad + \sum_{j=l-L_{MB}+1}^{l-1}\hat{\mathbf{h}}_{SB}^{[N_I]}(j)\cdot\hat{\mathbf{h}}_{SB}^{[N_I]}(j)^H\Bigg\}, \quad (49) \end{aligned}$$

where N_I 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 $\mathbf{C}_{\hat{\mathbf{h}}}^{[i]}(l)$.

$$\mathbf{U}\mathbf{D}\mathbf{V}^H = \mathtt{svd}\left(\mathbf{C}_{\hat{\mathbf{h}}}^{[i]}(l)\right), \quad (50)$$

$$\hat{r} = \arg\min_{r\leq W}\mathrm{MDL}\left\{\mathtt{diag}(\mathbf{D})|_{1:r}\right\}. \quad (51)$$

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-B* as shown in Fig. 2 when $N_{G1} = N_{G2} = 0$. It should be noted that we have to use the *input data range-B* because we can not estimate the length W CIR with the *input data range-A* as shown in Fig. 2, since $M_t^{EST} < W$ when $N_t = W$, where M_t^{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 $\mathbf{y}_t(l) \in \mathbb{C}^{N_t+W-1}$ can be described in the same way as that in (1), as:

$$\begin{aligned}\mathbf{y}_t(l) &= \mathbf{H}_t(l)\mathbf{s}_t(l) + \mathbf{H}'_t(l-1)\mathbf{s}'_t(l-1) \\ &\quad + \mathbf{H}''_t(l)\mathbf{s}''_t(l) + \mathbf{n}_t, \end{aligned} \tag{52}$$

where

$$\mathbf{s}_t(l) = \mathbf{x}_t(l) \in \mathbb{C}^{N_t}, \tag{53}$$

$$\mathbf{s}'_t(l-1) = \mathbf{x}_d(l-1)|_{(N_d-N_t+1):N_d} \in \mathbb{C}^{N_t}, \tag{54}$$

and

$$\mathbf{s}''_t(l) = \mathbf{x}_d(l)|_{1:N_t} \in \mathbb{C}^{N_t} \tag{55}$$

if $N_{CP} = 0$. Otherwise, $\mathbf{s}''_t(l)$ indicates the data portion including the CP in the current burst,[4] as

$$\mathbf{s}''_t(l) = \begin{bmatrix} \mathbf{x}_d(l)|_{(N_d-W+1):N_d} \\ \mathbf{x}_d(l)|_{1:(N_d-W)} \end{bmatrix} \in \mathbb{C}^{N_t}. \tag{56}$$

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 vectors $\mathbf{h}(l)$, $\mathbf{h}(l-1)$ and $\mathbf{h}(l)$, respectively. The noise vector $\mathbf{n}_t \in \mathbb{C}^{N_t+W-1}$ follows $\mathcal{CN}(0, \sigma_n^2)$.

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

$$\begin{aligned}\tilde{\mathbf{y}}_t^{[i]}(l) &= \mathbf{y}_t(l) - \Big\{\hat{\mathbf{H}}_t'^{[i-1]}(l-1)\,\hat{\mathbf{s}}_t'^{[i-1]}(l-1) \\ &\quad + \hat{\mathbf{H}}_t''^{[i-1]}(l)\,\hat{\mathbf{s}}_t''^{[i-1]}(l)\Big\}, \end{aligned} \tag{57}$$

where $\hat{\mathbf{H}}_t'^{[i-1]}(l-1)$, $\hat{\mathbf{s}}_t'^{[i-1]}(l-1)$, $\hat{\mathbf{H}}_t''^{[i-1]}(l)$ and $\hat{\mathbf{s}}_t''^{[i-1]}(l)$ are obtained as the result of the $(i-1)-$th iteration. We initialize $\hat{\mathbf{H}}_t'^{[i-1]}(l-1) = \hat{\mathbf{H}}_t''^{[i-1]}(l) = \mathbf{O}$ and $\hat{\mathbf{s}}_t'^{[i-1]}(l-1) = \hat{\mathbf{s}}_t''^{[i-1]}(l) = \mathbf{0}$ for the first iteration ($i = 1$). For the burst located at the head of the frame, we may exploit the result obtained at the final N_I-th iteration in the previous frame. That is, $\hat{\mathbf{H}}_t'^{[i-1]}(l-1) = \hat{\mathbf{H}}_t'^{[N_I]}(l-1)$ and $\hat{\mathbf{s}}_t'^{[i-1]}(l-1) = \hat{\mathbf{s}}_t'^{[N_I]}(l-1)$, for any i when $l = 1 + (f-1)N_B$ with f being the frame number. It is expected that CHATES with the single burst channel estimation (58) using (57), (59) and (60) improves the estimation accuracy:

$$\hat{\hat{\mathbf{h}}}_{SB}^{[i]}(l) = \tilde{\mathbf{R}}_{\mathbf{XX}}^{[i]}(l)^{-1}\,\tilde{\mathbf{R}}_{\mathbf{XY}}^{[i]}(l), \tag{58}$$

$$\tilde{\mathbf{R}}_{\mathbf{XX}}^{[i]}(l) = \tilde{\mathbf{X}}_t(l)^H\,\tilde{\mathbf{X}}_t(l) + \gamma^{[i-1]}(l)\,\hat{\mathbf{X}}_d^{[i-1]}(l)^H\,\hat{\mathbf{X}}_d^{[i-1]}(l), \tag{59}$$

$$\tilde{\mathbf{R}}_{\mathbf{XY}}^{[i]}(l) = \tilde{\mathbf{X}}_t(l)^H\,\tilde{\mathbf{y}}_t^{[i]}(l) + \gamma^{[i-1]}(l)\,\hat{\mathbf{X}}_d^{[i-1]}(l)^H\,\mathbf{y}_d(l), \tag{60}$$

where $\tilde{\mathbf{X}}_t(l) \in \mathbb{C}^{(N_t+W-1)\times W}$ is a Toeplitz matrix whose first column vector is $[\mathbf{x}_t^T(l),\ \mathbf{0}_{1\times W-1}]^T$. CHATES with multi-burst channel estimation can be also implemented with the length of training sequence $N_t = W$.

The CHATES differs from the conventional channel estimation techniques in the sense that CHATES performs IBI cancellation (57) while the conventional techniques [10] do not. It should be noted that the computational complexity order of CHATES is equivalent to that of the counterpart of the conventional technique. Because IBI cancellation (57) requires $O(W^2 + WN_t)$ which is less than the complexity order $O(N_dW^2)$ for the channel estimation part of (58). On the other hand, the conventional technique (45) also requires the complexity order $O(N_dW^2)$. Therefore, the computational order for (58) is, dominated by the channel estimation part, $O(N_dW^2)$ which is equivalent to that of (45), when $W = N_t \ll N_d$.

V. Simulations

This section describes results of computer simulations conducted to verify the feasibility and effectiveness of the proposed techniques. To make fair comparison, account is taken of the spectrum efficiency η of the structure of the burst format, with which the average SNR used in the simulations is connected to the average energy per bit to noise density ratio (E_b/N_0), as

$$\text{SNR} = \eta \cdot E_b/N_0, \tag{61}$$

$$\eta = R_c \cdot M_b \cdot \frac{N_d}{K}, \tag{62}$$

where the modulation multiplicity $M_b = 1$ for BPSK.

The parameters used in the following simulations are detailed 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

$$\mathbf{P_x} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \tag{63}$$

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.

[4] In the case of $N_d = W$, $\mathbf{s}''_t(l)$ is defined as the CP section only: $\mathbf{s}''_t(l) = \mathbf{x}_d(l)|_{(N_d-W+1):N_d}$.

TABLE I
BURST FORMATS.

Format No.	N_t	N_{G1}	N_{CP}	N_d	N_{G2}	R_c	η
1	64	0	0	256	0	1/2	0.4
2	64	0	64	192	0	2/3	0.4
3	64	64	0	256	64	1/2	0.29

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^e_{EQU} between the LLR λ^e_{EQU} (23) and the coded bits input to the symbol mapper c_M is defined by

$$\begin{aligned} \mathrm{I}^e_{EQU} &= \mathrm{I}(\lambda^e_{EQU}; c_M) \\ &= \frac{1}{2}\sum_{m=\pm 1}\int_{-\infty}^{+\infty} \Pr(\lambda^e_{EQU}|m)\log_2\frac{\Pr(\lambda^e_{EQU}|m)}{\Pr(\lambda^e_{EQU})}d\lambda^e_{EQU}, \end{aligned}$$

where $\Pr(\lambda^e_{EQU}|m)$ is the conditional probability density of λ^e_{EQU} 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^a_{EQU}. This is because of the noise enhancement described in Section III-D. In contrast, *CHATUE2* improves I^e_{EQU} and achieves almost the same point as that with TEQ-CP when $\mathrm{I}^a_{EQU} = 1$, although its left most point at $\mathrm{I}^a_{EQU} = 0$ 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 $\mathrm{I}^e_{DEC} = 1$ without intersection in the channel realization used and hence the MI between the *a posteriori* LLR of decoder λ^p_{DEC} 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) CHATUE algorithms allows 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 $\mathrm{I}^e_{DEC} = 1$ 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.

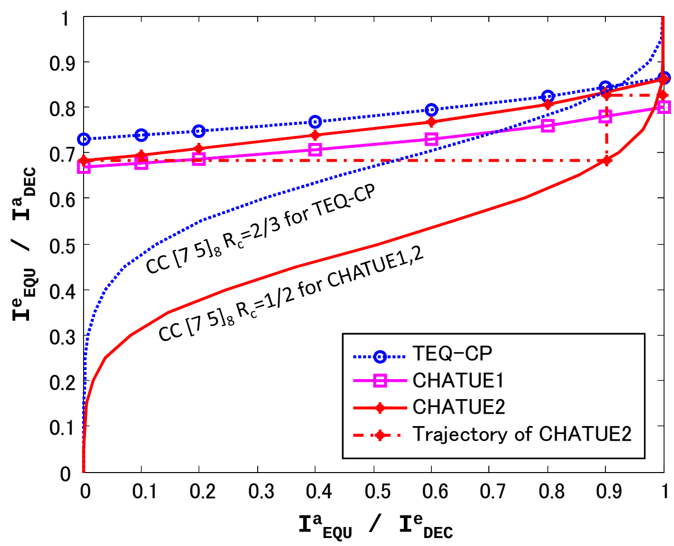

Fig. 3. EXIT charts and trajectory of iterative processing over a 64-path Rayleigh fading at average SNR = 2.4 dB.

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 dB $\leq$ SNR $\leq$ 16 dB which corresponds to 0 dB $\leq E_b/N_0 \leq$ 20 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 $L_{MB} \to \infty$ [10], given by

$$\mathrm{MSE}_{\mathrm{CRB}}(\sigma_n) = \frac{r}{M^{est}_t + M^{est}_d}\sigma^2_n, \tag{64}$$

where r is the number of dominant paths. M^{est}_t and M^{est}_d 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.

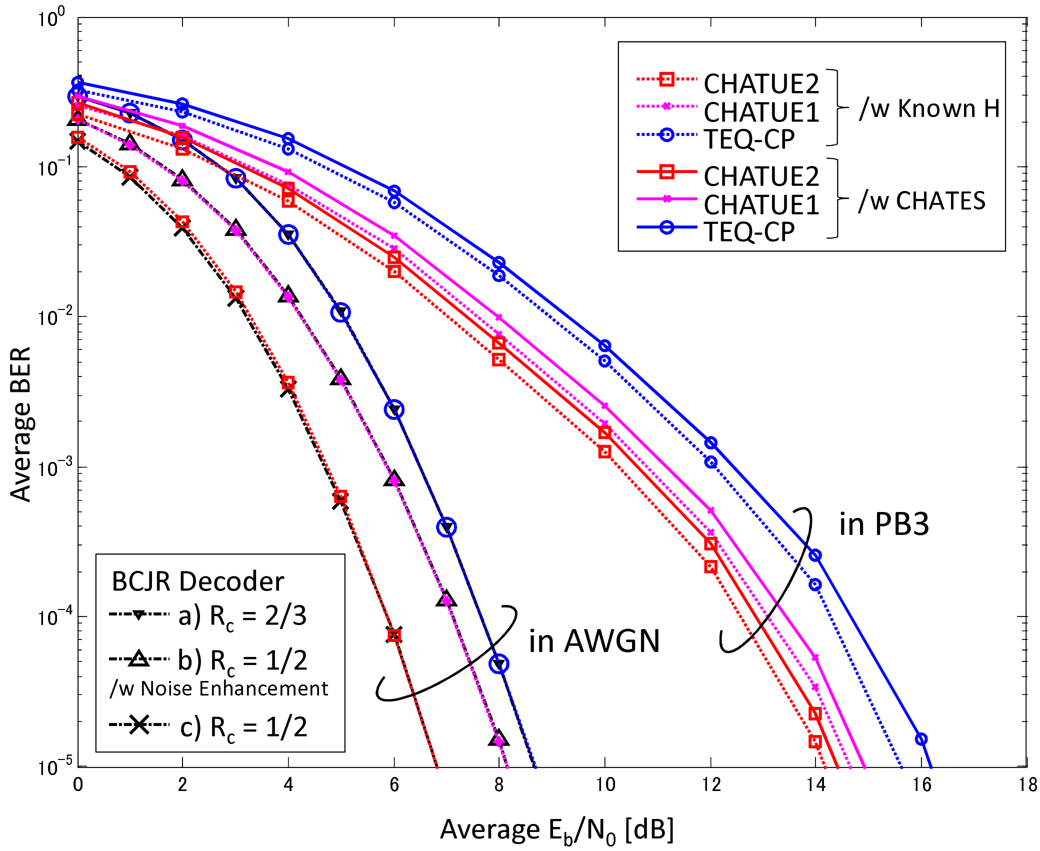

Fig. 5. BER for 1-path static AWGN vs. Average BER in PB3 with 6 iterations: The spectrum efficiency is fixed to $\eta = 0.4$. *CHATUE1* and *CHATUE2* use *Burst Format 1*, whereas TEQ-CP uses *Burst Format 2*.

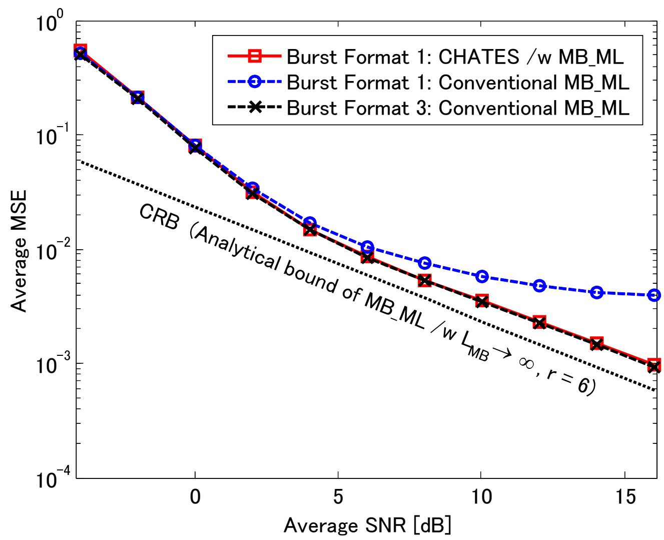

Fig. 4. Average MSE of $\hat{\mathbf{h}}$ in PB3 with 6 iterations.

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.[5] 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

[5]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 $2\sigma_n^2$ 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).

proposed *CHATUE2* can fully exploit the time duration made available by eliminating the CP, which allows for the use of a lower rate code $(R_c = 1/2)$ when the channel estimate is accurate enough.

Fig. 5 also shows the BER performance of turbo equalization for PB3 with CHATES using MB_ML and ideal channel estimation. *CHATUE2* with ideal channel estimation improves the BER over *CHATUE1* by 0.5 dB at BER = 10^{-5} since the proposed technique with the composite replica improves the SNR at the equalizer output. *CHATUE2* with ideal channel estimation achieves a gain of about 1.5 dB over TEQ-CP at BER = 10^{-5} because the CHATUE algorithms allows for the use of lower rate codes.

To verify the applicability of the proposed technique in realistic scenarios, we then present results of verification simulations conducted using channel sounding measurement data. The measurement campaign took place at the court yard of Technical University of Ilmenau in Germany. The RUSK channel sounder [13] was used for the measurement campaign. The channel impulse response data shows, as observed in [14], that the peak position varies quite frequently, which, does not happen with the model-based simulations such as the PB3 channel model. The channel obtained in the field measurement has up to 45 symbols of ISI. Average power control was assumed. The sounder's transmitter moved at a pedestrian speed. Fig. 6 shows the BER performance in this propagation scenario. The parameters related to the burst format and the receiver's algorithm were set at the same values as that in the verification in PB3. *CHATUE2* with ideal channel estimation achieves a gain of 0.5 dB over *CHATUE1*, and a gain of at least 1 dB than TEQ-CP at BER = 10^{-5}. With CHATES using MB_ML, *CHATUE2* outperforms *CHATUE1* and requires 1 dB lower E_b/N_0 to achieve BER = 10^{-5} than TEQ-CP. It should be noted that the approximation (31) does not cause any numerical instability, even with estimated CIR, as we identified in Figs. 5 and 6.

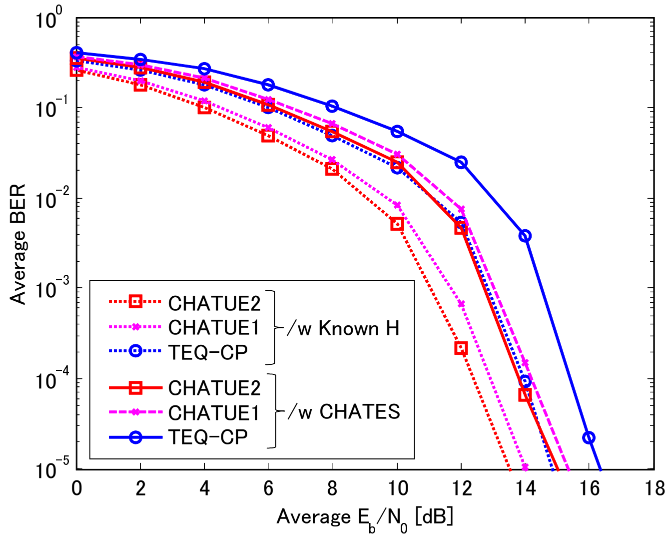

Fig. 6. Average BER using Measurement Data with 6 iterations: The spectrum efficiency is fixed to $\eta = 0.4$. CHATUE1 and CHATUE2 use *Burst Format 1*, whereas TEQ-CP uses *Burst Format 2*.

VI. Conclusions

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 (CHATES), that improves spectrum efficiency without sacrificing estimation accuracy. CHATES can be applied to *CHATUE1*, *CHATUE2* and TEQ-CP, although CHATES 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 CHATES 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 $\mathrm{E}[||\hat{\mathbf{s}}(l)||^2] \to 1$ after enough iterations, the noise covariance matrix (12) converges to

$$\mathbf{\Omega}(l) \quad \to \quad \sigma_n^2 \frac{N_d + L}{N_d} \mathbf{I}_{N_d}. \tag{65}$$

Hence, the equation (16) converges to

$$\mathbf{\Gamma}(l) \quad \to \quad \frac{N_d}{(N_d + L)\sigma_n^2} \mathbf{I}_{N_d}, \tag{66}$$

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

[1] K. Anwar, Z. Hui, and T. Matsumoto, “Chained turbo equalization for block transmission without guard interval,” in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, May. 2010, pp. 1 – 5.

[2] K. Kansanen, “Wireless broadband single-carrier systems with MMSE turbo equalization receivers,” *University of Oulu, Oulu Finland*, 2005.

[3] K. Anwar and T. Matsumoto, “Low complexity time-concatenated turbo equalization for block transmission without guard interval: Part 1 - the concept,” *Wireless Personal Comm., Springer, DOI 10.1007/s11277-012-0563-0*, Apr. 2012.

[4] Z. Hui, K. Anwar, and T. Matsumoto, “Low complexity time-concatenated turbo equalization for block transmission without guard interval: Part 2 - application to SC-FDMA,” *Wireless Personal Comm., Springer, DOI 10.1007/s11277-011-0409-1*, Sep. 2011.

[5] D. Wang, C. Wei, Z. Pan, X. You, C. H. Kyu, and J. B. Jang, “Low-complexity turbo equalization for single-carrier systems without cyclic prefix,” in *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*, Nov. 2008, pp. 1091 –1095.

[6] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate (Corresp.),” *Information Theory, IEEE Trans. on*, vol. 20, no. 2, pp. 284 – 287, Mar. 1974.

[7] K. Kansanen and T. Matsumoto, “An analytical method for MMSE MIMO turbo equalizer EXIT chart computation,” *Wireless Communications, IEEE Transactions on*, vol. 6, no. 1, pp. 59 –63, Jan. 2007.

[8] X. Wang and H. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *Communications, IEEE Transactions on*, vol. 47, no. 7, pp. 1046 –1061, Jul. 1999.

[9] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” *Communications, IEEE Transactions on*, vol. 49, no. 10, pp. 1727 –1737, Oct. 2001.

[10] M. Nicoli, S. Ferrara, and U. Spagnolini, “Soft-iterative channel estimation: Methods and performance analysis,” *Signal Processing, IEEE Trans. on*, vol. 55, no. 6, pp. 2993 –3006, Jun. 2007.

[11] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *Acoustics, Speech and Signal Processing, IEEE Trans. on*, vol. 33, no. 2, pp. 387 – 392, Apr. 1985.

[12] 3rd Generation Partnership Project, “3GPP TR25.996, Spatial channel model for MIMO simulations (Release 6),” 2003.

[13] http://www.channelsounder.de/.

[14] Y. Takano, K. Anwar, and T. Matsumoto, “Performance of turbo equalization using doped accumulator with channel estimation,” in *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*, Dec. 2011, pp. 1 –6.

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 (RWSf06), 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.

Tad Matsumoto (S’84.SM’95.F’10) received his BS, MS, and PhD degrees from Keio University, Yokohama, Japan, in 1978, 1980, and 1991, respectively, all in electrical engineering. He joined Nippon Telegraph and Telephone Corporation (NTT) in April 1980. Since he engaged in NTT, he was involved in a lot of research and development projects, all for mobile wireless communications systems. In July 1992, he transferred to NTT DoCoMo, where he researched Code-Division Multiple-Access techniques for Mobile Communication Systems. In April 1994, he transferred to NTT America, where he served as a Senior Technical Advisor of a joint project between NTT and NEXTEL Communications. In March 1996, he returned to NTT DoCoMo, where he served as a Head of the Radio Signal Processing Laboratory until August of 2001; He worked on adaptive signal processing, multipleinput multiple-output turbo signal detection, interference cancellation, and space-time coding techniques for broadband mobile communications. In March 2002, he moved to University of Oulu, Finland, where he served as a Professor at Centre for Wireless Communications. In 2006, he served as a Visiting Professor at Ilmenau University of Technology, Ilmenau, Germany, funded by the German MERCATOR Visiting Professorship Program. Since April 2007, he has been serving as a Professor at Japan Advanced Institute of Science and Technology (JAIST), Japan, while also keeping the position at University of Oulu. Prof. Matsumoto has been appointed as a Finland Distinguished Professor for a period from January 2008 to December 2012, funded by the Finnish National Technology Agency (Tekes) and Finnish Academy, under which he preserves the rights to participate in and apply to European and Finnish national projects. Prof. Matsumoto is a recipient of IEEE VTS Outstanding Service Award (2001), Nokia Foundation Visiting Fellow Scholarship Award (2002), IEEE Japan Council Award for Distinguished Service to the Society (2006), IEEE Vehicular Technology Society James R. Evans Avant Garde Award (2006), and Thuringen State Research Award for Advanced Applied Science (2006), 2007 Best Paper Award of Institute of Electrical, Communication, and Information Engineers of Japan (2008), Telecom System Technology Award by the Telecommunications Advancement Foundation (2009), IEEE Communication Letters Exemplifying Reviewer Award (2011) and the UK Royal Academy of Engineering Distinguished Visiting Fellowship Award (2012). He is a Fellow of IEEE and a Member of IEICE. He is serving as an IEEE Vehicular Technology Distinguished Lecturer during the term July 2011 – June 2013.