

PAPER

Adaptive MAP Detection via the EM Algorithm for LDPC-Coded MIMO-OFDM Mobile Communications

Tsuyoshi KASHIMA^{†a)}, Kazuhiko FUKAWA^{†b)}, *Members*, and Hiroshi SUZUKI^{†c)}, *Fellow*

SUMMARY This paper proposes an iterative maximum *a posteriori* probability (MAP) receiver for multiple-input-multiple-output (MIMO) and orthogonal frequency-division multiplexing (OFDM) mobile communications. For exploiting the space, time, and frequency diversity, the low-density parity-check code (LDPC) is used as a channel coding with a built-in interleaver. The receiver employs the expectation maximization (EM) algorithm so as to perform the MAP symbol detection with reasonable computational complexity. The minimum mean square error (MMSE), recursive least squares (RLS), and least mean square (LMS) algorithms are theoretically derived for the channel estimation within this framework. Furthermore, the proposed receiver performs a new scheme called backward symbol detection (BSD), in which the signal detection uses the channel impulse response that is estimated one OFDM symbol later. The advantage of BSD, which is explained from the viewpoint of the message passing algorithm, is that BSD can exploit information on the both precedent and subsequent OFDM symbols, similarly to RLS with *smoothing* and *removing* (SR-RLS) [25]. In comparison with SR-RLS, BSD reduces the complexity at the cost of packet error rate (PER) performance. Computer simulations show that the receiver employing RLS for the channel estimation outperforms the ones employing MMSE or LMS, and that BSD can improve the PER performance of the ones employing RLS or LMS.

key words: mobile communication, MIMO, OFDM, LDPC, MAP, EM algorithm, channel estimation

1. Introduction

For future high bit-rate mobile communications, orthogonal frequency-division multiplexing (OFDM) and multiple-input-multiple-output (MIMO) are a promising set of transmission techniques. OFDM is a multi-carrier transmission scheme with a guard interval (GI), and can maintain excellent transmission performance even in multi-path fading channels. MIMO is a spatial multiplexing and diversity scheme using multiple antennas, and can increase the spectral efficiency.

Coding techniques such as the space-time code [3], the space-frequency code [4], [5], and the space-time-frequency code [6] have been studied for integrating the spatial diversity and channel coding for MIMO-OFDM. However, they assume time or frequency coherence of a channel, which is not always satisfied in mobile radio communications. Conversely, a channel code with an interleaver in the time, frequency, and space domains randomizes the coherence, and

can exploit diversity effects in the three domains. Since the low-density parity-check (LDPC) code [1], [2] has a built-in interleaver and great ability in error-correction, LDPC-coded MIMO-OFDM systems are expected to achieve better performance than the conventional spatial coding techniques [17], [18].

The optimum receiver for the coded systems performs signal detection on the maximum *a posteriori* (MAP) criterion. Since the MAP detection requires a prohibitively large amount of computational complexity, several low-complexity algorithms have been proposed. As the low-complexity versions, the interference canceller [7]–[9] and the equalizers [10], [11] perform joint processing of decoding and symbol detection by using *a priori* information on coded bits. Another low-complexity receiver employs the expectation-maximization (EM) algorithm [12], [13], and performs joint processing of symbol detection and channel estimation in order to improve accuracy of the channel estimation [14]–[18].

As for the LDPC-coded MIMO-OFDM system, the receiver proposed in [18] applies the EM algorithm to the MAP detection. This receiver assumes a time-invariant channel model and employs the minimum mean square error (MMSE) algorithm for the channel estimation. In contrast, this paper proposes a receiver possessing the recursive least squares (RLS) and least mean square (LMS) channel estimators. The proposed receiver shows greater ability in fading tracking because it assumes a time-variant channel model. Within the framework of the EM algorithm, we derive the MMSE, RLS and LMS channel estimators according to channel models.

In addition, this paper proposes a joint processing method of the recursive channel estimation and the symbol detection in the iteration of the EM algorithm. The method is referred to as backward symbol detection (BSD), in which the signal detection uses the channel impulse response that is estimated one OFDM symbol later. Although our conference paper [24] has empirically found BSD, this paper theoretically discusses BSD from the viewpoint of the factor graph [19], [20]. The theoretical analysis shows that BSD indirectly uses information on all OFDM symbols through iterations, without direct effect of the targeted OFDM symbol. Therefore, BSD gains the similar effect to RLS with *smoothing* and *removing* (SR-RLS) [25] and can improve the packet error ratio (PER) performance. Note that BSD is inferior in the PER performance but superior in complexity to SR-RLS. It is also noteworthy that the proposed scheme

Manuscript received November 14, 2005.

Manuscript revised July 18, 2006.

[†]The authors are with Tokyo Institute of Technology, Tokyo, 152-8550 Japan.

a) E-mail: tsuyoshi.kashima@nokia.com

b) E-mail: fukawa@radio.ss.titech.ac.jp

c) E-mail: suzuki@radio.ss.titech.ac.jp

DOI: 10.1093/ietcom/e90-b.2.312

can be applied to a system using other channel codes such as the Turbo code if the channel decoders can provide *a priori* information.

The rest of this paper is organized as follows. Section 2 presents the system model in both time and frequency representation. Section 3 describes the iterative joint MAP receiver with the LDPC decoding and the EM algorithm. Section 4 explains the joint processing of the MAP symbol detection and discusses BSD. In Sect. 5, the MMSE, LMS and RLS algorithms are derived within the framework of the EM algorithm. Section 6 shows computer simulation results. Finally, some concluding remarks are given in Sect. 7.

2. System Model

Let us consider an LDPC-coded MIMO-OFDM system with N_T transmitter antennas, N_R receiver antennas, N subcarriers, and N_S OFDM symbols in one packet.

A block diagram of the transmitter is shown in Fig. 1. An LDPC-encoded bit sequence is modulated into a serial symbol sequence, and the serial-parallel transform converts it into $N_T N$ parallel symbol sequences. Let $b_k(i, n)$ denote the transmitted symbol from the k -th ($1 \leq k \leq N_T$) transmitter antenna at the n -th ($0 \leq n < N$) subcarrier in the i -th ($0 \leq i < N_S$) OFDM symbol. Each N parallel symbol sequences is passed into the inverse fast Fourier transform (IFFT) processor corresponding to each transmitter antenna, and GI is added to its output every OFDM symbol. Let Δ_t be the OFDM sampling interval, that is the inverse of the bandwidth. With Δ_t as a time unit, the GI length is given by $\Delta_G \Delta_t$. Thus, the transmitted symbol $s_k(m)$ from the k -th transmitter antenna at discrete time $m\Delta_t$ can be expressed as

$$s_k(m) = \sum_{n=0}^{N-1} b_k(i, n) \exp \left\{ j \frac{2\pi n [m - (i+1)\Delta_G]}{N} \right\}, \quad (1)$$

where $i(N + \Delta_G) \leq m < (i+1)(N + \Delta_G)$.

Let us assume that every channel between transmitter and receiver antennas is subject to uncorrelated multipath Rayleigh fading and is also spatially uncorrelated to each other. $h_{kl}(t, \tau)$ represents the channel impulse response between the k -th transmitter antenna and the l -th ($1 \leq l \leq N_R$) receiver antenna, and is a function of delay time τ at time t . $h_{kl}(t, \tau)$ can be modeled as

$$h_{kl}(t, \tau) = \sum_{d=0}^D \alpha_{kl,d}(t) \delta(\tau - d\Delta_t), \quad (2)$$

where $D\Delta_t$ is the maximum delay time, $\alpha_{kl,d}(t)$ represents the Rayleigh-fading complex envelope on the d -th ($0 \leq d \leq D$) propagation path between the k -th transmitter antenna and the l -th receiver antenna at time t , and $\delta(\cdot)$ is the Dirac delta function.

A block diagram of the receiver is shown in Fig. 2. Let $y_l(m)$ be the signal received by the l -th receiver antenna at discrete time $m\Delta_t$, and $y_l(m)$ is given by

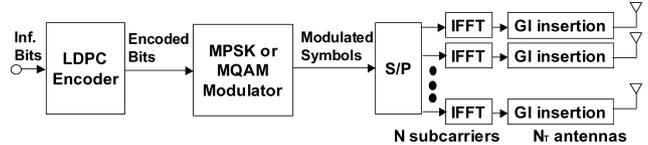


Fig. 1 LDPC-coded MIMO-OFDM transmitter.

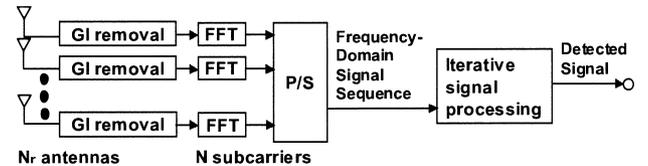


Fig. 2 MIMO-OFDM receiver.

$$y_l(m) = \sum_{k=1}^{N_T} \sum_{d=0}^D \alpha_{lk,d}(m\Delta_t) s_k(m-d) + n_l(m), \quad (3)$$

where $n_l(m)$ is the noise of the l -th receiver antenna at discrete time $m\Delta_t$. GI is removed from the received signal every OFDM symbol, and the resultant signal is fed into the fast Fourier transform (FFT) processor. $Y_l(i, n)$ denotes the FFT output of the l -th receiver antenna at the n -th subcarrier in the i -th OFDM symbol, and can be expressed as

$$Y_l(i, n) = N^{-1} \sum_{m=i(N+\Delta_G)+\Delta_G}^{(i+1)(N+\Delta_G)-1} y_l(m) \times \exp \left\{ -j \frac{2\pi n [m - (i+1)\Delta_G]}{N} \right\}. \quad (4)$$

On the assumption that the maximum delay time is not longer than GI, that is $D \leq \Delta_G$, and that $\alpha_{lk,d}(m\Delta_t)$ is constant during one OFDM symbol, substituting (1) and (3) into (4) approximately yields

$$Y_l(i, n) = \sum_{k=1}^{N_T} H_{lk}(i, n) b_k(i, n) + N_l(i, n), \quad (5)$$

$$H_{lk}(i, n) = \sum_{d=0}^D \alpha_{lk,d}[m_a(i)\Delta_t] \exp \left(-j \frac{2\pi n d}{N} \right), \quad (6)$$

$$N_l(i, n) = N^{-1} \sum_{m=i(N+\Delta_G)+\Delta_G}^{(i+1)(N+\Delta_G)-1} n_l(m) \exp \left\{ -j \frac{2\pi n [m - (i+1)\Delta_G]}{N} \right\} \quad (7)$$

where $\alpha_{lk,d}[m_a(i)\Delta_t]$ is an average value of $\alpha_{lk,d}(m\Delta_t)$ in the i -th OFDM symbol, and $m_a(i) = i(N + \Delta_G) + \Delta_G + N/2$ is assumed in this paper.

The symbol detection uses the frequency-domain signal model of (5) in the vector form as

$$\begin{aligned} Y_l(i, n) &= \mathbf{H}_l^H(i, n) \mathbf{X}(i, n) + N_l(i, n), \\ \mathbf{H}_l^H(i, n) &= [H_{l1}(i, n), \dots, H_{lN_T}(i, n)], \\ \mathbf{X}^H(i, n) &= [b_1^*(i, n), \dots, b_{N_T}^*(i, n)], \end{aligned} \quad (8)$$

where $\mathbf{H}_l(i, n)$ and $\mathbf{X}(i, n)$ are N_T -by-1 vectors, and the superscript H and the asterisk denote Hermitian transposition and complex conjugation, respectively. Furthermore, an N -by-1 vector $\mathbf{Y}_l(i)$ having $Y_l(i, n)$ as its elements is given by

$$\begin{aligned} \mathbf{Y}_l(i) &= \mathbf{X}^{\text{H}}(i)\mathbf{H}_l(i) + \mathbf{N}(i), \\ \mathbf{X}^{\text{H}}(i) &= [\mathbf{X}_1(i), \dots, \mathbf{X}_{N_T}(i)], \\ \mathbf{X}_k(i) &= \text{diag}[b_k(i, 1), \dots, b_k(i, N)], \\ \mathbf{H}_l^{\text{H}}(i) &= [H_{l1}^*(i, 1), \dots, H_{l1}^*(i, N), \dots, \\ &\quad H_{lN_T}^*(i, 1), \dots, H_{lN_T}^*(i, N)], \end{aligned} \quad (9)$$

where $\mathbf{N}(i)$ is an N -by-1 vector having $N_l(i, n)$ as its elements, $\mathbf{H}_l(i)$ is a NN_T -by-1 vector, $\mathbf{X}(i)$ is a NN_T -by- N matrix, and $\text{diag}[\]$ denotes the diagonal matrix consisting of its arguments.

On the other hand, the channel estimation of $\mathbf{H}_l(i)$ uses the time-domain signal model of (3) in the vector form as

$$\begin{aligned} y_l(m) &= \mathbf{h}_l^{\text{H}}(m)\mathbf{x}(m) + n_l(m), \\ \mathbf{h}_l^{\text{H}}(m) &= [\alpha_{l1,0}(m\Delta_t), \dots, \alpha_{l1,D}(m\Delta_t), \dots \\ &\quad \alpha_{lN_T,0}(m\Delta_t), \dots, \alpha_{lN_T,D}(m\Delta_t)], \\ \mathbf{x}^{\text{H}}(m) &= [s_1^*(m), \dots, s_1^*(m-D), \dots \\ &\quad s_{N_T}^*(m), \dots, s_{N_T}^*(m-D)], \end{aligned} \quad (10)$$

where $\mathbf{h}_l(m)$ and $\mathbf{x}(m)$ are $(D+1)N_T$ -by-1 vectors.

Two major advantages of this time-domain estimation over the frequency-domain one, which directly estimates $\mathbf{H}_l(i)$, are as follows. One is that the time-domain one updates the estimate every Δ_t and can track fast fading more rapidly than the frequency-domain one, of which the update period is one OFDM symbol $(N + \Delta_G)\Delta_t$. The other is that the latter needs additional information on correlation of frequency responses among subcarriers in order to maintain the same accuracy as that of the former.

For simplicity, the receiver antenna index l will be omitted below.

3. Iterative MAP Receiver

3.1 Receiver Structure

A block diagram of the proposed MAP receiver is illustrated in Fig. 3. The receiver performs the following two steps. At the first step, the maximum likelihood (ML) detector produces the log likelihood ratio (LLR) $\lambda_1(p)$ of every p -th coded bit. Then the LDPC decoder performs the LDPC decoding once by using $\lambda_1(p)$ and generates *a posteriori* LLR

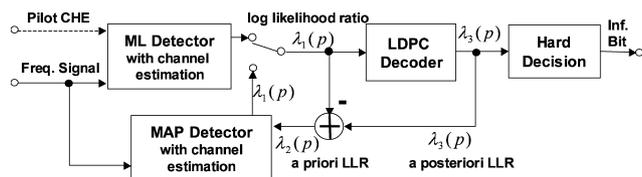


Fig. 3 Iterative MAP receiver.

$\lambda_3(p)$ of the coded bit. At the second step, the MAP detector operates instead of the ML detector, and iterates the MAP-based signal detection every LDPC decoding. By using *a priori* LLR $\lambda_2(p)$ of the coded bit, which is given by $\lambda_2(p) = \lambda_3(p) - \lambda_1(p)$ from the decoder, the MAP detector updates $\lambda_1(p)$. To achieve better performance, the MAP detector iterates MAP demodulation and channel estimation within the EM algorithm scheme. The iteration of the MAP detector and the LDPC decoder is referred to as the turbo iteration, and repeats unless the LDPC-parity check is passed or the iteration number exceeds a predetermined number. The iteration of the MAP demodulation and the channel estimation is referred to as the EM iteration, and repeats unless the iteration number exceeds another predetermined number.

3.2 MAP Criterion for Packet Transmission

The packet is assumed to consist of a preamble part \mathbf{X}_p and a data part \mathbf{X}_d . Let \mathbf{Y}_p and \mathbf{Y}_d represent the received signals corresponding to \mathbf{X}_p and \mathbf{X}_d . Since \mathbf{X}_p is known and \mathbf{X}_d is to be detected, the MAP detection should maximize the probability $P_{APP} = P(\mathbf{X}_d | \mathbf{X}_p, \mathbf{Y}_d, \mathbf{Y}_p)$ that can be rewritten as

$$P_{APP} = P(\mathbf{X}_d, \mathbf{X}_p | \mathbf{Y}_d, \mathbf{Y}_p) \frac{P(\mathbf{Y}_d, \mathbf{Y}_p)}{P(\mathbf{X}_p, \mathbf{Y}_d, \mathbf{Y}_p)}. \quad (11)$$

The second factor does not depend on \mathbf{X}_d , and the maximization of P_{APP} is equivalent to that of the first factor. Thus, the MAP detection can be performed by maximizing $P(\mathbf{X} | \mathbf{Y})$ where \mathbf{X} and \mathbf{Y} denote the sets of $(\mathbf{X}_d, \mathbf{X}_p)$ and $(\mathbf{Y}_d, \mathbf{Y}_p)$, respectively. From now on, we assume that the packet has one OFDM preamble symbol followed by LDPC-coded data OFDM symbols.

3.3 MAP Decoding of LDPC

The decoding of LDPC uses the sum-product algorithm [1], [2]. The calculation is performed in the logarithm domain to reduce the computational complexity. Figure 4 shows a schematic diagram of the Tanner graph, which represents the LDPC code structure. Variable nodes v_p ($p = 1 \dots P$) and check nodes c_q ($q = 1 \dots Q$) represent the coded bits and the parity check relations, respectively. Let \mathbf{C} denote the Q -by- P LDPC parity check matrix, and $C_{qp} = 1$ only if a connection between v_p and c_q exists in the Tanner graph.

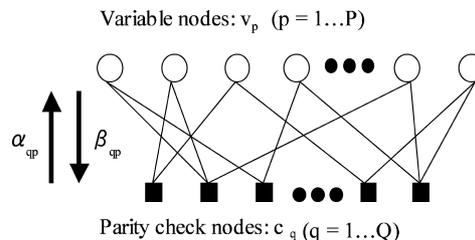


Fig. 4 Schematic diagram of LDPC code structure.

Also, let α_{qp} and β_{qp} denote *extrinsic* LLR from c_q to v_p , and *a priori* LLR from v_p to c_q , respectively. After the initialization that $\beta_{qp} = 0$ for all p, q that satisfy $\mathbf{C}_{qp} = 1$, the decoding iteration including the update of α_{qp} and β_{qp} starts. By using the LLR $\lambda_1(p)$ from the demodulators and β_{qp} , α_{qp} is updated through the parity check nodes as

$$\alpha_{qp} = \left[\prod_{p' \in V(q) \setminus p} \text{sign} \{ \lambda_1(p') + \beta_{qp'} \} \right] \times f \left[\sum_{p' \in V(q) \setminus p} f \{ | \lambda_1(p') + \beta_{qp'} | \} \right], \quad (12)$$

where

$$f(x) = \ln \frac{\exp(x) + 1}{\exp(x) - 1}, \quad (13)$$

$$\text{sign}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}. \quad (14)$$

$V(q)$ is a set of variable nodes connected to the check node c_q , and $\setminus p$ means that v_p is excluded from the set. On the other hand, β_{qp} is updated using α_{qp} through the variable nodes as

$$\beta_{qp} = \sum_{q' \in C(p) \setminus q} \alpha_{q'p}, \quad (15)$$

where $C(p)$ is a set of check nodes connected to the variable node v_p . Then, a temporary *a posteriori* LLR is given by

$$\lambda_3(p) = \lambda_1(p) + \sum_{q' \in C(p)} \alpha_{q'p}. \quad (16)$$

Every iteration, the LDPC decoder calculates the hard estimation of coded bits as

$$\hat{w}(p) = \begin{cases} 0, & \lambda_3(p) \geq 0 \\ 1, & \lambda_3(p) < 0 \end{cases}. \quad (17)$$

For every turbo iteration, the LDPC decoding (12)–(17) operates only once and checks whether $[\hat{w}(1), \dots, \hat{w}(P)]\mathbf{C}^t = 0$ is satisfied or not. Here, the superscript t denotes transposition. If the turbo iteration continues, *a posteriori* LLR $\lambda_3(p)$ of (16) is fed into the MAP detector for the next iteration.

3.4 MAP Detection via EM Algorithm

The MAP detection is performed in the frequency domain and maximizes $P(\mathbf{X}|\mathbf{Y})$ as

$$\hat{\mathbf{X}} = \arg \max_{\mathbf{X}} \log P(\mathbf{X}|\mathbf{Y}). \quad (18)$$

This operation requires a prohibitively large amount of computational complexity, because it needs to consider all possible cases of the unobservable channel impulse response. The EM algorithm can asymptotically find the solution of (18) with reasonable complexity, by iterating the following two steps [12], [13]:

$$\text{E step : } Q(\mathbf{X}|\mathbf{X}^{(r)}) = E_{\mathbf{H}} \left[\log P(\mathbf{Y}|\mathbf{H}, \mathbf{X}) | \mathbf{Y}, \mathbf{X}^{(r)} \right], \quad (19)$$

$$\text{M step : } \mathbf{X}^{(r+1)} = \arg \max_{\mathbf{X}} \left[Q(\mathbf{X}|\mathbf{X}^{(r)}) + \log P(\mathbf{X}) \right], \quad (20)$$

where \mathbf{H} denotes the channel impulse response of the whole packet, and $E_{\mathbf{H}}\{\cdot | \mathbf{Y}, \mathbf{X}^{(r)}\}$ is the expectation operation with respect to \mathbf{H} following the distribution of $P(\mathbf{H}|\mathbf{Y}, \mathbf{X}^{(r)})$. $\mathbf{X}^{(r)}$ is the estimate of \mathbf{X} at the r -th EM iteration. Note that the operation of (19) and (20) always satisfies $P(\mathbf{X}^{(r+1)}|\mathbf{Y}) \geq P(\mathbf{X}^{(r)}|\mathbf{Y})$ [12], [13].

Since the noise vector $\mathbf{N}(i)$ in (9) is statistically independent with respect to index i , $\log P(\mathbf{Y}|\mathbf{H}, \mathbf{X})$ can be decoupled into terms of individual OFDM symbols, and $E_{\mathbf{H}}\{\cdot\}$ for the i -th OFDM symbol is equivalent to $E_{\mathbf{H}(i)}\{\cdot\}$. Therefore, (19) can be calculated as

$$Q(\mathbf{X}|\mathbf{X}^{(r)}) = \sum_{i=1}^{N_s} E_{\mathbf{H}(i)} \left[\frac{\|\mathbf{Y}(i) - \mathbf{X}^{\mathbf{H}}(i)\mathbf{H}(i)\|^2}{\sigma^2} | \mathbf{Y}, \mathbf{X}^{(r)} \right] + c, \quad (21)$$

where c is a term independent of \mathbf{X} . $E\{\mathbf{N}(i)\mathbf{N}^{\mathbf{H}}(i)\} = \sigma^2 \mathbf{I}_N$ was used for the derivation, where \mathbf{I}_N is an N -by- N identity matrix and σ^2 is the average noise power in frequency domain. Each term in the sum of (21) can be rewritten as

$$\begin{aligned} & E_{\mathbf{H}(i)} \left[\|\mathbf{Y}(i) - \mathbf{X}^{\mathbf{H}}(i)\mathbf{H}(i)\|^2 | \mathbf{Y}, \mathbf{X}^{(r)} \right] \\ &= E_{\mathbf{H}(i)} \left[\left\| \mathbf{Y}(i) - \mathbf{X}^{\mathbf{H}}(i)\hat{\mathbf{H}}(i) + \mathbf{X}^{\mathbf{H}}(i) \{ \hat{\mathbf{H}}(i) - \mathbf{H}(i) \} \right\|^2 \right], \\ &= \|\mathbf{Y}(i) - \mathbf{X}^{\mathbf{H}}(i)\hat{\mathbf{H}}(i)\|^2 + \text{tr} \left[\mathbf{X}^{\mathbf{H}}(i)\boldsymbol{\Sigma}_{\mathbf{H}}(i)\mathbf{X}(i) \right], \end{aligned} \quad (22)$$

where tr denotes the trace operation and the property of the trace was used for the derivation. $\hat{\mathbf{H}}(i)$ and $\boldsymbol{\Sigma}_{\mathbf{H}}(i)$ are the mean and covariance matrix of $\mathbf{H}(i)$, which are defined as

$$\begin{aligned} \hat{\mathbf{H}}(i) &= E_{\mathbf{H}} \left[\mathbf{H}(i) | \mathbf{Y}, \mathbf{X}^{(r)} \right], \\ \boldsymbol{\Sigma}_{\mathbf{H}}(i) &= E_{\mathbf{H}(i)} \left[(\mathbf{H}(i) - \hat{\mathbf{H}}(i))(\mathbf{H}(i) - \hat{\mathbf{H}}(i))^{\mathbf{H}} | \mathbf{Y}, \mathbf{X}^{(r)} \right]. \end{aligned} \quad (23)$$

Furthermore, it is evident from the definitions of $\mathbf{Y}(i)$ and $\mathbf{X}(i)$ that the norm and the trace in (22) can be decoupled into terms of individual subcarriers. Thus, from (21) and (22), $Q(\mathbf{X}|\mathbf{X}^{(r)})$ can be rewritten as

$$Q(\mathbf{X}|\mathbf{X}^{(r)}) = \sum_{i=1}^{N_s} \sum_{n=1}^N q[\mathbf{X}(i, n) | \mathbf{X}^{(r)}] + c, \quad (25)$$

and then the E-step of (19) becomes equivalent to

$$\begin{aligned} \text{E step : } & q[\mathbf{X}(i, n) | \mathbf{X}^{(r)}] = \\ & \frac{|Y(i, n) - \hat{\mathbf{H}}^{\mathbf{H}}(i, n)\mathbf{X}(i, n)|^2}{\sigma^2} - \frac{\mathbf{X}^{\mathbf{H}}(i, n)\boldsymbol{\Sigma}_{\mathbf{H}}(i, n)\mathbf{X}(i, n)}{\sigma^2} \end{aligned} \quad (26)$$

where $\hat{\mathbf{H}}(i, n)$ and $\boldsymbol{\Sigma}_{\mathbf{H}}(i, n)$ are the mean and N_T -by- N_T covariance matrix of $\mathbf{H}(i, n)$ defined as

$$\hat{\mathbf{H}}(i, n) = E_{\mathbf{H}} [\mathbf{H}(i, n) | \mathbf{Y}, \mathbf{X}^{(r)}], \quad (27)$$

$$\Sigma_{\mathbf{H}}(i, n) = \begin{bmatrix} [\Sigma_{\mathbf{H}}(i)]_{(n,n)} & \cdots & [\Sigma_{\mathbf{H}}(i)]_{(n,T+n)} \\ \vdots & \ddots & \vdots \\ [\Sigma_{\mathbf{H}}(i)]_{(T+n,n)} & \cdots & [\Sigma_{\mathbf{H}}(i)]_{(T+n,T+n)} \end{bmatrix} \quad (28)$$

with $T = (N_T - 1)N$.

Furthermore, on the assumption that the correlation between different transmitted subcarrier symbols is negligible, $\log P(\mathbf{X})$ of (20) also can be decoupled into terms of individual subcarriers. Then, the M-step of (20) becomes equivalent to

$$\begin{aligned} \text{M step : } \mathbf{X}^{(r+1)}(i, n) = \\ \arg \max_{\mathbf{X}(i,n)} \{q[\mathbf{X}(i, n) | \mathbf{X}^{(r)}] + \log P[\mathbf{X}(i, n)]\}. \end{aligned} \quad (29)$$

Consequently, the MAP demodulation part of the EM algorithm results in (26) and (29). Note that different from the symbol detection, the channel estimation cannot be decoupled because it needs to utilize the correlation between subcarriers.

4. Joint Processing of Channel Estimation and Symbol Detection

4.1 Several Types of Joint Processing

This subsection discusses the relation between the symbol detection of $\mathbf{X}(i)$ and the channel estimation of $\mathbf{H}(i)$. Figures 5 and 6 show the relation in the cases of time-invariant and time-variant channel models, respectively. In these figures, the solid and dotted arrows indicate which information is used for the channel estimation and the symbol detection, respectively. With the time-invariant model, the channel estimation uses all OFDM symbols directly and corresponds to MMSE. With the time-variant model, the channel estimation recursively updates estimates every OFDM symbol and

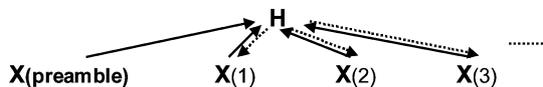


Fig. 5 Message passing for time-invariant channel model.

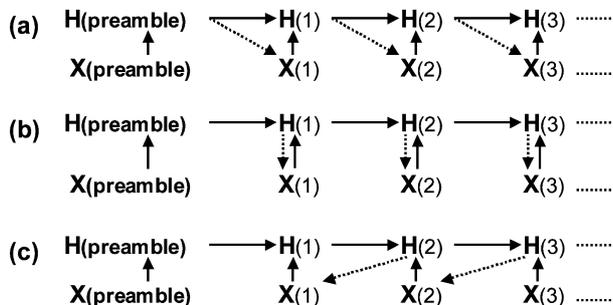


Fig. 6 Message passing for time-variant channel model. (a) Forward (FSD), (b) Normal (NSD), (c) Backward (BSD).

corresponds to the RLS and LMS.

Let us discuss three methods of the symbol detection with the same recursive channel estimation, which are depicted by (a)–(c) in Fig. 6. Forward symbol detection (FSD), normal symbol detection (NSD), and backward symbol detection (BSD) use $[\hat{\mathbf{H}}(i - 1, n), \Sigma_{\mathbf{H}}(i - 1, n)]$, $[\hat{\mathbf{H}}(i, n), \Sigma_{\mathbf{H}}(i, n)]$, and $[\hat{\mathbf{H}}(i + 1, n), \Sigma_{\mathbf{H}}(i + 1, n)]$ for the symbol detection of $\mathbf{X}(i, n)$ in (26), respectively. At the first EM iteration, the detection and the estimation need to be alternately processed from the preamble, and the MAP receiver can conduct only FSD. However, FSD uses only information on the precedent symbols and cannot exploit results of the estimation and detection of the previous EM iteration. Thus, NSD or BSD should be employed at the second or later EM iterations because they can exploit results of the previous EM iteration.

4.2 Message Passing Algorithm on Factor Graph

To consider NSD and BSD from the viewpoint of the message-passing algorithm on the factor graph, we briefly explain this algorithm [19], [20]. Figure 7 shows the factor graph representation of the system with the time-variant channel model. The factor graph is a bipartite graph consisting of *local functions* and their arguments called *variables*. The *variables* correspond to the information on both the channel impulse response and the transmitted symbols, whereas the *local functions* correspond to the relations between them. White circles and black squares symbolize the *variables* and the *local functions*, respectively. The advantage of the message-passing algorithm on the factor graphs is that it can easily calculate global functions such as the probability of related variables, by only using local message passing in a unified manner. According to the rule of the message passing, three types of local message passing are derived as shown in Fig. 8. They are the original recursive channel estimation, backward recursive channel estimation for *smoothing*, and *removing* for symbol detection. *Smoothing* modifies the RLS algorithm so that the recursive update is performed both forward and backward in order to utilize all symbols for the channel estimation. *Removing* modifies the RLS algorithm in order to perform channel estimation without using the targeted OFDM symbol directly. The RLS algorithm with *smoothing* and *removing* is referred to as SR-RLS, and improves the PER performance [25].

The EM algorithm performs the channel estimation on the condition of \mathbf{Y} and $\mathbf{X}^{(r)}$, and thus should utilize the information on all OFDM symbols for the channel estimation.

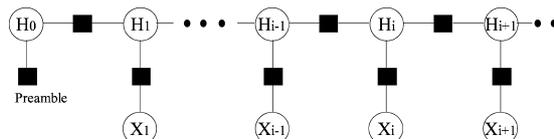


Fig. 7 Factor graph representation of system with time-variant channel model.

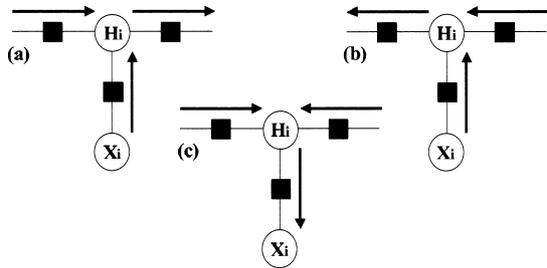


Fig. 8 Local propagation on the factor graph. (a) forward recursive channel estimation, (b) backward recursive channel estimation for *smoothing*, (c) *removing* for symbol detection.

Smoothing modifies the RLS algorithm to satisfy this requirement. With the direct feedback between the channel estimation and the symbol detection, however, the channel estimation error due to the present noise can degrade the signal detection, and vice versa, which is called noise amplification. Hence, *removing* is also important because it can alleviate this performance degradation.

4.3 Comparison between BSD and SR-RLS

SR-RLS outperforms the original RLS channel estimation that uses only information on the precedent symbols. However, the computational complexity of SR-RLS is about three times larger than that of RLS [25]. Conversely, BSD and NSD require the same amount of complexity as the original RLS because they do not change the estimation algorithm at all in contrast to SR-RLS. BSD and NSD differ in only that BSD performs the signal detection by using the impulse response that is estimated one OFDM symbol later. This modification enables BSD to exploit information on the subsequent symbols in addition to the precedent symbols, which achieves the similar effect to *smoothing*. It also can prevent the channel estimation from using the present symbol directly, which has the similar effect as *removing*. Thus, BSD is expected to achieve better performance than NSD without any additional complexity.

Consequently, BSD is superior to SR-RLS in the computational complexity, and also has the similar effect to *smoothing* and *removing*, which can avoid the severe noise amplification problem. Detailed performance comparisons require simulations, which are explained in Sect. 6.

5. Channel Estimation

This section derives MMSE, RLS, and LMS for the channel estimation within the framework of the EM algorithm. In the following calculations, the sets of the transmitted symbols and the received signals in the time domain are denoted by

$$\mathcal{X}(m) = \{\mathbf{x}(0), \dots, \mathbf{x}(m)\}, \quad \mathcal{X} = \mathcal{X}(N_m - 1), \quad (30)$$

$$\mathcal{Y}(m) = \{\mathbf{y}(0), \dots, \mathbf{y}(m)\}, \quad \mathcal{Y} = \mathcal{Y}(N_m - 1), \quad (31)$$

where $N_m = N_S(N + \Delta_G)$ is the number of sampling points in one packet.

Note that the channel estimation estimates the mean $\hat{\mathbf{h}}(m)$ and the covariance matrix $\mathbf{P}(m)$ of $\mathbf{h}(m)$ where σ_n^2 denotes the average power of $n(m)$ and $\mathbf{P}(m)$ is normalized by σ_n^2 for notation simplicity. By applying the discrete Fourier transform, $\hat{\mathbf{H}}(i)$ and $\Sigma_{\mathbf{H}}(i)$ are calculated as $\hat{\mathbf{H}}(i) = \mathbf{W}\hat{\mathbf{h}}^*(m_a)$ and $\Sigma_{\mathbf{H}}(i) = \sigma_n^2 \mathbf{W}\mathbf{P}^*(m_a)\mathbf{W}^H$, where \mathbf{v}^* generally denotes the vector, each element of which is the complex conjugation of the correspondent element of \mathbf{v} , the NN_T -by- $(D + 1)N_T$ vector \mathbf{W} is defined as $\mathbf{W} = \text{diag}[\mathbf{w}, \dots, \mathbf{w}]$, and the N -by- $(D + 1)$ vector \mathbf{w} is given by $[\mathbf{w}]_{n+1,d+1} = \exp(-j2\pi nd/N)$.

5.1 MMSE

The MMSE channel estimation assumes the time-invariant channel model. Thus, $\mathbf{h}(m)$, $\hat{\mathbf{h}}(m)$, and $\mathbf{P}(m)$ are rewritten as \mathbf{h} , $\hat{\mathbf{h}}$, and \mathbf{P} , respectively. \mathbf{h} satisfies

$$P(\mathbf{h}|\mathcal{X}, \mathcal{Y}) = P(\mathcal{Y}|\mathbf{h}, \mathcal{X})P(\mathbf{h})/P(\mathcal{Y}|\mathcal{X}). \quad (32)$$

On the assumption that \mathbf{h} and \mathcal{Y} are subject to multivariate complex Gaussian distributions, $P(\mathbf{h}|\mathcal{X}, \mathcal{Y})$, $P(\mathcal{Y}|\mathbf{h}, \mathcal{X})$, and $P(\mathbf{h})$ are given by

$$P(\mathbf{h}|\mathcal{X}, \mathcal{Y}) \propto \exp\left[-(\mathbf{h} - \hat{\mathbf{h}})^H \mathbf{P}^{-1}(\mathbf{h} - \hat{\mathbf{h}})/\sigma_n^2\right], \quad (33)$$

$$P(\mathcal{Y}|\mathbf{h}, \mathcal{X}) \propto \prod_{m=0}^{N_m-1} \exp\left[-\|y(m) - \mathbf{h}^H \mathbf{x}(m)\|^2/\sigma_n^2\right], \quad (34)$$

$$P(\mathbf{h}) \propto \exp\left(-\mathbf{h}^H \mathbf{P}_0^{-1} \mathbf{h}/\sigma_n^2\right), \quad (35)$$

where $\mathbf{P}_0 = \text{diag}[\sigma_{1,0}^2, \dots, \sigma_{1,D}^2, \dots, \sigma_{N_T,0}^2, \dots, \sigma_{N_T,D}^2]$ is the delay profile, and $\sigma_{k,d}^2$ is the average power of the d -th propagation path from the k -th transmitter antenna, which are assumed to be known to the receiver. By substituting (33)–(35) into (32), taking the logarithm, and comparing the terms of the first and second degree of \mathbf{h} , the following equations are derived.

$$\mathbf{h}^H \mathbf{P}^{-1} \mathbf{h} = \sum_{m=0}^{N_m-1} \mathbf{h}^H \mathbf{x}(m) \mathbf{x}^H(m) \mathbf{h} + \mathbf{h}^H \mathbf{P}_0^{-1} \mathbf{h}, \quad (36)$$

$$\mathbf{h}^H \mathbf{P}^{-1} \hat{\mathbf{h}} = \sum_{m=0}^{N_m-1} \mathbf{h}^H \mathbf{x}(m) y^*(m). \quad (37)$$

The necessary and sufficient condition for any \mathbf{h} to satisfy these equations, which is referred to as the MMSE algorithm, is given by

$$\hat{\mathbf{h}} = \mathbf{R}^{-1} \mathbf{V}, \quad \mathbf{P} = \mathbf{R}^{-1}, \quad (38)$$

$$\mathbf{R} = \sum_{m=0}^{N_m-1} \mathbf{x}(m) \mathbf{x}^H(m) + \mathbf{P}_0^{-1}, \quad \mathbf{V} = \sum_{m=0}^{N_m-1} \mathbf{x}(m) y^*(m). \quad (39)$$

5.2 RLS

The RLS algorithm assumes the time-variant channel model, and has much greater ability in channel tracking than MMSE. In general, $\mathbf{h}(m)$ satisfies

$$P[\mathbf{h}(m)|\mathcal{Y}(m), \mathcal{X}] = \frac{P[\mathbf{h}(m)|\mathcal{Y}(m-1), \mathcal{X}] P[y(m)|\mathbf{h}(m), \mathcal{X}]}{P[y(m)|\mathcal{Y}(m-1), \mathcal{X}]} \quad (40)$$

Since $\mathbf{h}(m)$ is assumed to be a multivariate complex Gaussian process,

$$P[\mathbf{h}(m)|\mathcal{Y}(m-1), \mathcal{X}] \propto \exp\left(-\Delta_{m|m-1}^H \mathbf{P}_{m|m-1}^{-1} \Delta_{m|m-1} / \sigma_n^2\right), \quad (41)$$

$$P(\mathbf{h}(m)|\mathcal{Y}(m), \mathcal{X}) \propto \exp\left(-\Delta_{m|m}^H \mathbf{P}_{m|m}^{-1} \Delta_{m|m} / \sigma_n^2\right), \quad (42)$$

$$P(y(m)|\mathbf{h}(m), \mathcal{X}) \propto \exp\left[-\|y(m) - \mathbf{h}^H(m)\mathbf{x}(m)\|^2 / \sigma_n^2\right] \quad (43)$$

hold, where $\Delta_{m|m'} = \mathbf{h}(m) - \hat{\mathbf{h}}_{m|m'}$, and $\hat{\mathbf{h}}_{m|m'}$ and $\mathbf{P}_{m|m'}$ are the mean and the normalized covariance of $\mathbf{h}(m)$, respectively, on the condition of $\mathcal{Y}(m')$ and $\mathcal{X}(m')$. By substituting (41)–(43) into (40), taking the logarithm, and comparing the terms of the first and second degree of $\mathbf{h}(m)$, the following equations are derived.

$$\mathbf{h}^H(m) \mathbf{P}_{m|m-1}^{-1} \mathbf{h}(m) + \mathbf{h}^H(m) \mathbf{x}(m) \mathbf{x}^H(m) \mathbf{h}(m) = \mathbf{h}^H(m) \mathbf{P}_{m|m}^{-1} \mathbf{h}(m), \quad (44)$$

$$\mathbf{h}^H(m) \mathbf{P}_{m|m-1}^{-1} \hat{\mathbf{h}}_{m|m-1} + \mathbf{h}^H(m) \mathbf{x}(m) y^*(m) = \mathbf{h}^H(m) \mathbf{P}_{m|m}^{-1} \hat{\mathbf{h}}_{m|m}. \quad (45)$$

The necessary and sufficient condition for any $\mathbf{h}(m)$ to satisfy these equations is given by

$$\mathbf{P}_{m|m-1}^{-1} + \mathbf{x}(m) \mathbf{x}^H(m) = \mathbf{P}_{m|m}^{-1}, \quad (46)$$

$$\mathbf{P}_{m|m-1}^{-1} \hat{\mathbf{h}}_{m|m-1} + \mathbf{x}(m) y^*(m) = \mathbf{P}_{m|m}^{-1} \hat{\mathbf{h}}_{m|m}. \quad (47)$$

Then, substituting $\mathbf{P}_{m|m-1}^{-1}$ of (46) into (47) obtains

$$\hat{\mathbf{h}}_{m|m} = \hat{\mathbf{h}}_{m|m-1} + \mathbf{K}(m) [y_m^* - \mathbf{x}^H(m) \hat{\mathbf{h}}_{m|m-1}], \quad (48)$$

$$\mathbf{K}(m) = \mathbf{P}_{m|m} \mathbf{x}(m). \quad (49)$$

Applying the matrix inversion lemma [23] to (46) yields

$$\mathbf{P}_{m|m} = \mathbf{P}_{m|m-1} - \frac{\mathbf{P}_{m|m-1} \mathbf{x}(m) \mathbf{x}^H(m) \mathbf{P}_{m|m-1}}{\mathbf{x}^H(m) \mathbf{P}_{m|m-1} \mathbf{x}(m) + 1}. \quad (50)$$

Substituting (50) into (49) results in

$$\mathbf{K}(m) = \mathbf{P}_{m|m-1} \mathbf{x}(m) [\mathbf{x}^H(m) \mathbf{P}_{m|m-1} \mathbf{x}(m) + 1]^{-1}. \quad (51)$$

By using (51), (50) can be rewritten as

$$\mathbf{P}_{m|m} = \mathbf{P}_{m|m-1} - \mathbf{K}(m) \mathbf{x}^H(m) \mathbf{P}_{m|m-1}. \quad (52)$$

Furthermore, to obtain the relations between $\hat{\mathbf{h}}_{m|m-1}$ and $\hat{\mathbf{h}}_{m-1|m-1}$, and between $\mathbf{P}_{m|m-1}$ and $\mathbf{P}_{m-1|m-1}$, $\mathbf{h}(m)$ is assumed to be subject to a random walk model as

$$\mathbf{h}(m) = \mathbf{h}(m-1) + \mathbf{q}(m-1), \quad (53)$$

where $\mathbf{q}(m-1)$ is a complex vector, the elements of which

are white Gaussian processes with zero mean and independent of $\mathbf{h}(m-1)$. For simplicity, its covariance matrix is assumed to be $\langle \mathbf{q}^H(m-1) \mathbf{q}(m-1) \rangle = q \mathbf{P}_{m-1|m-1}$ where $\langle \cdot \rangle$ is the averaging operation. Hence, the following equations are derived from (53)

$$\hat{\mathbf{h}}_{m|m-1} = \hat{\mathbf{h}}_{m-1|m-1}, \quad (54)$$

$$\mathbf{P}_{m|m-1} = \lambda^{-1} \mathbf{P}_{m-1|m-1}, \quad (55)$$

where λ is the forgetting factor given by $\lambda = 1/(1+q)$, and takes the value from 0 to 1.

Substituting (54) and (55) into (48), (51), and (52), and letting $\hat{\mathbf{h}}(m) = \hat{\mathbf{h}}_{m|m}$ and $\mathbf{P}(m) = \mathbf{P}_{m|m}$ result in the RLS algorithm as

$$\mathbf{K}(m) = \mathbf{P}(m-1) \mathbf{x}(m) [\mathbf{x}^H(m) \mathbf{P}(m-1) \mathbf{x}(m) + \lambda]^{-1}, \quad (56)$$

$$\hat{\mathbf{h}}(m) = \hat{\mathbf{h}}(m-1) + \mathbf{K}(m) [y^*(m) - \mathbf{x}^H(m) \hat{\mathbf{h}}(m-1)], \quad (57)$$

$$\mathbf{P}(m) = \lambda^{-1} [\mathbf{P}(m-1) - \mathbf{K}(m) \mathbf{x}^H(m) \mathbf{P}(m-1)]. \quad (58)$$

To obtain $\hat{\mathbf{h}}(m)$ and $\mathbf{P}(m)$, this recursive form should start from $m=0$, which is in the preamble.

5.3 LMS

The LMS algorithm can be derived from the RLS algorithm by assuming that $\mathbf{P}(m) = \mu \mathbf{I}_{(D+1)N_T}$ where μ is the step size of the LMS algorithm. Since $\mathbf{K}(m) = \mu \mathbf{x}(m)$ is obtained from (49), (57) can be rewritten as

$$\hat{\mathbf{h}}(m) = \hat{\mathbf{h}}(m-1) + \mu \mathbf{x}(m) [y^*(m) - \mathbf{x}^H(m) \hat{\mathbf{h}}(m-1)]. \quad (59)$$

Owing to the simplification, LMS has much lower computational complexity than RLS, although it still has a higher fading tracking ability than MMSE.

5.4 Computational Complexity

Table 1 compares the computational complexity of the channel estimation algorithms. The numbers in the table cells are the complexity per one receiver antenna during one packet, which the algorithms with the $N_T D$ taps over N_S OFDM sampling points require. Since NSD and BSD differ only in the timing of the symbol detection, they do not make any difference in the computational complexity. The computational complexity of LMS is much lower than the others due to its simplification. The processing time of RLS and LMS can be lower than that of MMSE depending on their implementation, because they operate recursively. Furthermore, compared to the SR-RLS, RLS with BSD achieves lower complexity because it needs no additional processing, such as *smoothing* and *removing*.

Table 1 Computational complexity.

	complex addition	complex multiplication	square root
MMSE	$\frac{5}{6}(N_T D)^3 + N_m(N_T D)^2$	$\frac{5}{6}(N_T D)^3 + N_m(N_T D)^2$	$N_T D$
LMS with NSD	$2N_m(N_T D)$	$2N_m(N_T D)$	0
LMS with BSD	$2N_m(N_T D)$	$2N_m(N_T D)$	0
RLS with NSD	$2N_m(N_T D)^2$	$3N_m(N_T D)^2$	0
RLS with BSD	$2N_m(N_T D)^2$	$3N_m(N_T D)^2$	0
SR-RLS	$6N_m(N_T D)^2$	$6N_m(N_T D)^2$	0

6. Computer Simulation

6.1 Simulation Condition

The simulation conditions are summarized in Table 2. A 9-path Rayleigh channel with exponential decay was assumed. The complex envelope on each propagation path is statistically independent and based on the Jakes model [22]. It was also assumed that the maximum delay time is not longer than GI, and the maximum delay considered for the channel estimation was set equally to GI, $16\Delta_t$. The maximum Doppler frequency f_D was normalized by the OFDM symbol duration T_s . The forgetting factor of the RLS algorithm was set equally to 0.99 in order to keep the numerical stability of the algorithm. Since the RLS algorithm in the time domain performs the recursive updates every Δ_t , the forgetting factor of 0.99 is still proper to track the fading fluctuation. On the other hand, the step size of the LMS algorithm was set equally to 0.3, which does not theoretically correspond to the forgetting factor of 0.99. The step size of 0.3, however, is proper in spite of this discrepancy because the convergence speed of the LMS algorithm is slower than RLS. The numbers of turbo and EM iterations were set to 10 and 4 except Figs. 13 and 14. These values will be justified later. The LDPC code length and information bits are 4096 bits and 2051 bits, respectively. A randomly generated matrix with the column weight of 2 and the row weight of 4 is converted into a full-rank matrix via the Gaussian elimination, and the resultant is used as the LDPC parity check matrix. Thus, the coding rate is not exactly 1/2. Note that the difference of the LDPC code structure does not affect the findings of this paper due to the randomness in the generation process

In the following graphs, “turboMAP” and “ML+LDPC (no turbo)” indicate with the turbo iteration and without it, respectively. It was assumed that the receiver in the cases of “perfect CSI” and “MMSE” knows all channel information and the delay profile.

6.2 Effect of Doppler Frequency

Figures 9 and 10 show the effect of the Doppler frequency on the PER performance of the proposed MAP receiver with average bit energy to noise density ratio (E_b/N_0) of 8 and 20 dB, respectively. In both cases, RLS and LMS achieve lower PER than MMSE, and their gains become larger when

Table 2 Simulation condition.

MIMO antennas ($N_T \times N_R$)	2×2
Modulation	QPSK (Gray mapping [21]), 64 subcarriers OFDM
GI/OFDM symbol duration	$16\Delta_t/80\Delta_t$
Preamble length	1 OFDM symbol
Data length	16 OFDM symbols
Channel estimation	16 taps
RLS forgetting factor	0.99
LMS step size	0.3
Channel model	No spatial correlation, Independent 9-path Rayleigh with 0.8 exponential decay for each impulse response between antennas
EM iterations	4, otherwise mentioned
Turbo iterations	10, otherwise mentioned
Information bits	2051 bits
LDPC code length	4096 bits

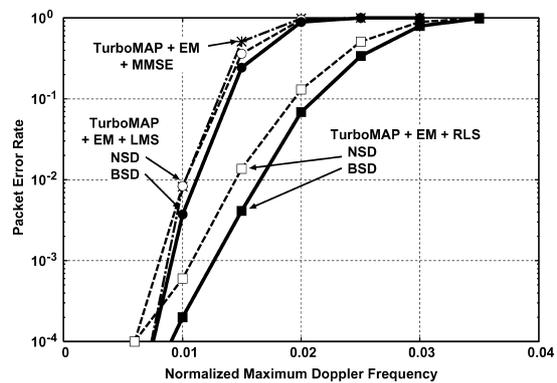


Fig. 9 Effect of Doppler frequency with E_b/N_0 of 8 dB.

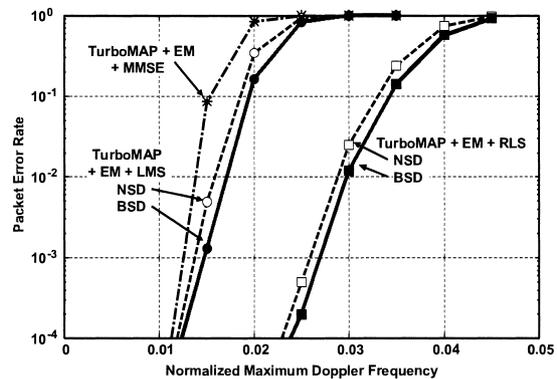


Fig. 10 Effect of Doppler frequency with E_b/N_0 of 20 dB.

the E_b/N_0 increases as the fading tracking ability becomes dominant. Furthermore, with both the LMS and RLS algorithms, the proposed BSD gains better PER performance than the conventional NSD. Especially, the improvement using RLS with E_b/N_0 of 8 dB is remarkable. This is because BSD avoids the noise amplification as explained in Sect. 4.2, and it is effective especially in low E_b/N_0 region. Also note that the improvement of BSD compared to NSD is available irrespectively of the Doppler frequency, and there is no

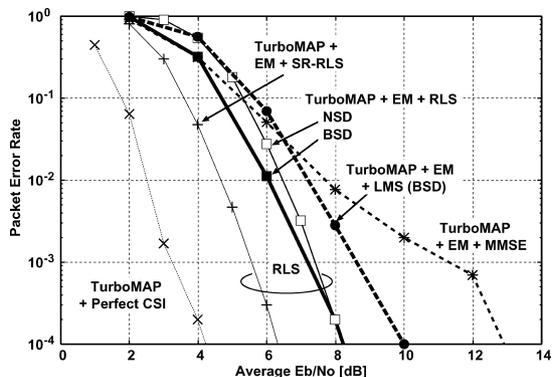


Fig. 11 Comparison of channel estimations ($f_D T_s = 0.01$).

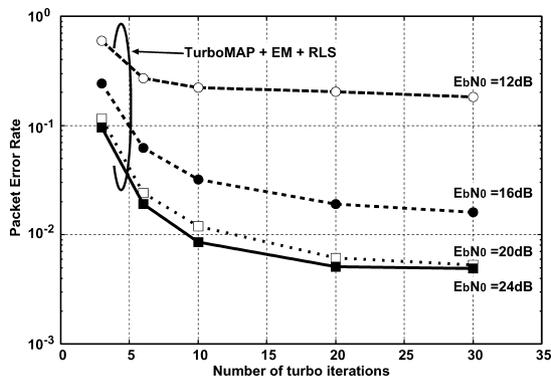


Fig. 13 Convergence in terms of the number of turbo iteration.

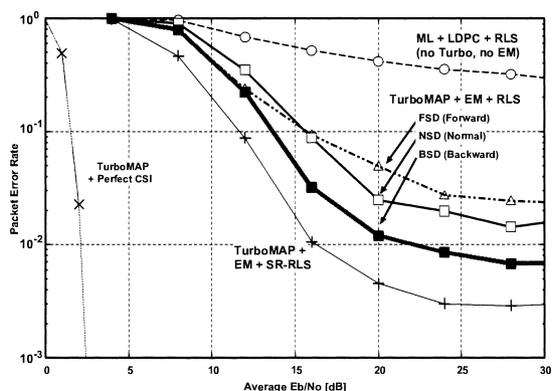


Fig. 12 Effect of turbo and EM iterations ($f_D T_s = 0.03$).

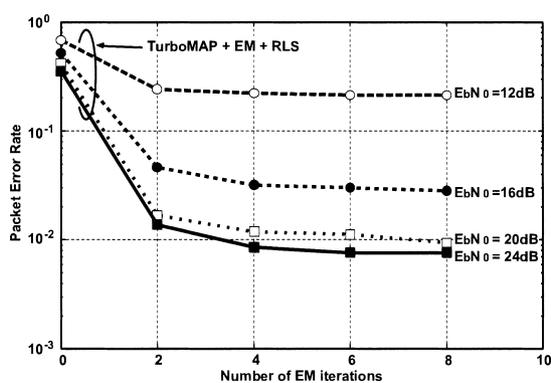


Fig. 14 Convergence in terms of the number of EM iteration.

threshold value where their PER performance crosses when $PER \geq 10^{-3}$.

6.3 Comparison between BSD and SR-RLS

Figure 11 shows PER of the proposed receivers with $f_D T_s = 0.01$. In spite of its better fading tracking ability, LMS with BSD is inferior to MMSE with average $E_b/N_0 < 6.5$ dB because LMS with the step size of 0.3 has much less ability in noise averaging than MMSE. LMS, however, maintains superiority to MMSE in computational complexity. RLS with BSD outperforms both MMSE and LMS with BSD irrespective of E_b/N_0 . RLS with NSD, however, is inferior to MMSE with average $E_b/N_0 < 5.5$ dB because NSD can utilize only the precedent OFDM symbols for the estimation, and cannot suppress the noise amplification.

Figure 12 shows PER performance with $f_D T_s = 0.03$. Since the MMSE and LMS channel estimators cannot track the fast fading, only the results with the RLS channel estimator are plotted in the figure. The turbo MAP iteration reduces PER by one or two orders with average $E_b/N_0 > 20$ dB. Since FSD in the EM iteration does not update the detected symbols, it corresponds to the receiver without the EM iteration. The PER performance of NSD is worse than that of FSD in low E_b/N_0 region due to the noise amplification problem. BSD achieves the best PER performance within FSD, NSD and BSD.

With both $f_D T_s = 0.01$ and 0.03 , RLS with BSD achieves the best performance among MMSE, LMS and RLS irrespective of E_b/N_0 . Although SR-RLS achieves better performance than RLS with BSD, RLS with BSD has three times lower computational complexity than SR-RLS.

6.4 EM and Turbo Iterations

Figure 13 shows the convergence speed of the turbo MAP iterations using RLS with BSD. Although 30 turbo iterations can achieve better PER than 10 iterations, it cannot reduce the remaining PER even to half. Figure 14 shows the convergence speed of the EM iterations of the same receiver. The improvement of PER by the first and second EM iterations is considerable, while more than four EM iterations improve little PER performance. From these simulation results, 10 turbo iterations and 4 EM iterations have been chosen for other simulations in this paper. Setting 7 or 8 turbo iterations and 2 EM iterations still has meaningful improvement. The choice of the numbers of iterations should consider the trade-off between the target performance and computational complexity.

7. Conclusion

This paper has proposed an iterative MAP receiver for LDPC-coded MIMO-OFDM mobile communications,

which employs the EM algorithm for the MAP symbol detection. Within the framework of the EM algorithm, the MMSE, RLS and LMS algorithms have been derived for the channel estimation according to different channel models. The proposed receiver performs a new joint processing of the symbol detection and the recursive channel estimation, which is referred to as BSD. In BSD, the symbol detection uses the channel impulse response that is estimated one OFDM symbol later. The advantage of BSD, which is explained from the viewpoint of the message-passing algorithm, is that BSD can exploit information on all but the targeted OFDM symbol, similarly to SR-RLS. In comparison with SR-RLS, BSD reduces the complexity at the cost of PER performance degradation.

Computer simulations showed that BSD improves the performance of both the RLS and LMS channel estimators, and that RLS outperforms either the MMSE or LMS algorithms. LMS is superior to MMSE in terms of its simplicity and better fading tracking ability.

References

- [1] R.G. Gallager, "Low density parity check codes," *IRE Trans. Inform. Theory*, vol.8, pp.21–28, Jan. 1962.
- [2] D.J.C. MacKay and R.M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Lett.*, vol.32, no.18, pp.1645–1646, Aug. 1996.
- [3] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband," *Proc. IEEE VTC*, vol.3, pp.2232–2236, Ottawa, Ont., Canada, May 1998.
- [4] K.F. Lee and D.B. Williams, "A space-frequency transmitter diversity technique for OFDM systems," *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol.3, pp.1473–1477, Nov. 2000.
- [5] H. Bolcskei and A. Paulraj, "Space-frequency coded broadband OFDM systems," *Proc. IEEE Wireless Commun. Networking Conf. (WCNC)*, vol.1, pp.1–6, Sept. 2000.
- [6] Z. Liu, Y. Xin, and G.B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Process.*, vol.50, no.10, pp.2465–2476, Oct. 2002.
- [7] I. Shon, "Space frequency block coded turbo-BLAST detection for MIMO-OFDM systems," *Electron. Lett.*, vol.39, no.21, pp.1557–1558, Oct. 2003.
- [8] J. Zhang, J. Olivier, A. Sayeed, and B.V. Veen, "Low complexity MIMO receiver via maximum SINR interference cancellation," *IEEE Veh. Technol. Conf. (VTC)*, vol.4, pp.2028–2032, May 2002.
- [9] S. Suyama, H. Suzuki, and K. Fukawa, "A MIMO-OFDM receiver employing the low-complexity turbo equalization in multipath environments with delay difference greater than the guard interval," *IEICE Trans. Commun.*, vol.E88-B, no.1, pp.39–46, Jan. 2005.
- [10] H.N. Lee and V. Gulati, "Iterative equalization/decoding of LDPC code transmitted over MIMO fading ISI channels," *Personal, Indoor and Mobile Radio Communications, IEEE International Symposium*, vol.3, pp.1330–1336, Sept. 2002.
- [11] A.M. Tonello, "MIMO MAP equalization and turbo decoding in interleaved space-time coded systems," *IEEE Trans. Commun.*, vol.51, no.2, pp.155–160, Feb. 2003.
- [12] G.J. McLachlan and T. Krishnan, *The EM algorithm and extensions*, John Wiley & Sons, 2000.
- [13] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J.R. Stat. Soc., Series B*, vol.39, pp.1–38, 1977.
- [14] B. Han, X. Gao, X. You, and M. Weckerle, "Joint channel estimation and symbol detection for SFBC-OFDM systems via the EM algorithm," *Commun.*, 2004 *IEEE International Conf.*, vol.6, pp.3148–3152, June 2004.
- [15] D.K.C. So and R.S. Cheng, "Improved iterative EM receiver for space time coded systems in frequency selective fading channel gain and order estimation," *IEEE VTC 2003 Spring*, vol.2, pp.847–851, April 2003.
- [16] S.Y. Park and C.G. Kang, "Complexity-reduced iterative MAP receiver for interference suppression in OFDM-based spatial multiplexing systems," *IEEE Trans. Veh. Technol.*, vol.53, no.5, pp.1316–1326, Sept. 2004.
- [17] B. Lu, G. Yue, and X. Wang, "Performance analysis and design optimization of LDPC-coded MIMO OFDM systems," *IEEE Trans. Signal Process.*, vol.52, no.2, pp.348–361, Feb. 2004.
- [18] B. Lu, X. Wang, and K.R. Narayanan, "LDPC-based space-time coded OFDM systems over correlated fading channels: Performance analysis and receiver design," *IEEE Trans. Commun.*, vol.50, no.1, pp.74–88, Jan. 2002.
- [19] F.R. Kschischang, B.J. Frey, and H.A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol.47, no.2, pp.498–519, Feb. 2001.
- [20] H.A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol.21, no.1, pp.28–41, Jan. 2004.
- [21] J. Wu and H.N. Lee, "Best mapping for LDPC coded modulation on SISO, MIMO and MAC channels," *IEEE Wireless Commun. Networking Conf.*, pp.2428–2431, March 2004.
- [22] P. Dent, G.E. Bottomley, and T.E. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol.29, pp.1162–1163, June 1993.
- [23] S. Haykin, *Adaptive Filter Theory*, 3rd ed., Prentice-Hall, 1996.
- [24] T. Kashima, K. Fukawa, and H. Suzuki, "Iterative-MAP adaptive detection via the EM algorithm for LDPC-coded MIMO-OFDM mobile communications in fast fading channel," *Proc. IEEE VTC 2005-Spring*, vol.3, pp.1681–1685, Stockholm, May/June 2005.
- [25] T. Kashima, K. Fukawa, and H. Suzuki, "Adaptive MAP receiver via the EM algorithm and message passing for MIMO-OFDM mobile communications," *IEEE J. Sel. Areas Commun.*, vol.24, no.3, pp.437–447, March 2006.



Tsuyoshi Kashima received the B.S. and M.S. degrees in theoretical physics from Tokyo University, Tokyo, Japan, in 1998, and 2000 respectively. Since joining Nokia Research Center, Tokyo in 2000, he was engaged in system level research on cellular and multi-hop networks, and is now engaged mainly in PHY and MAC cross-layer design for future radio systems. Since 2004, he has also started Ph.D. study in Tokyo Institute of Technology, Japan with continuing his work in Nokia. His Ph.D.

research topics are on the optimal receiver architecture utilizing iterative joint processing. Message-passing algorithm and Monte Carlo method are especially in his interest. Mr. Kashima is a member of IEEE.



Kazuhiko Fukawa received the B.S. and M.S. degrees in physics, and the Dr. Eng. degree in electrical and electronics engineering, all from Tokyo Institute of Technology, Tokyo, Japan, in 1985, 1987, and 1999 respectively. He joined Nippon Telegraph and Telephone Corporation (NTT), Japan, in 1987. Since then, he has been engaged in research on digital mobile radio communication systems and applications of the adaptive signal processing, including adaptive equalization, interference cancellation, and

adaptive arrays. He was a Senior Research Engineer at NTT Mobile Communications Network Inc. (NTT DoCoMo), from 1994 to 2000. Since April 2000, he has been an Associate Professor at the Tokyo Institute of Technology. Dr. Fukawa is a member of IEEE. He received the Paper Award from IEICE in 1995.



Hiroshi Suzuki received the B.S. degree in electrical engineering, the M.S. degree in physical electronics, and the Dr. Eng. Degree in electrical and electronics engineering, all from the Tokyo Institute of Technology, Tokyo, in 1972, 1974, and 1986, respectively. He joined the Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Japan, in 1974. He was engaged in research on devices in millimeter-wave regions. Since 1978, he has been engaged in fundamental and developmental researches on digital mobile communication systems.

He was an Executive Research Engineer in the Research and Development Department, NTT Mobile Communications Network, Inc. (NTT DoCoMo) from 1992 to 1996. Since September 1996, he has been a Professor at the Tokyo Institute of Technology. He is currently interested in various applications of the adaptive signal processing to radio signal transmission: adaptive arrays, multiuser detection, interference canceling, and MOMO-OFDM for future advanced multiple access communication systems. Dr. Suzuki is a member of IEEE. He received the Paper Award from the IEICE in 1995, and the award of IEICE Fellow in 2006.