Iterative MAP Channel Estimation Based on Factor Graph for OFDM Mobile Communications

Kazushi Muraoka, Kazuhiko Fukawa, Hiroshi Suzuki and Satoshi Suyama
Tokyo Institute of Technology
2-12-1, O-okayama, Meguro-ku, Tokyo, 152-8550 Japan
Email: {muraoka, fukawa, suzuki, ssuyama}@radio.ss.titech.ac.jp

***Abstract*—To maintain good transmission performance of orthogonal frequency-division multiplexing (OFDM) mobile communications over fast-fading channels, iterative maximum *a posteriori* (MAP) receivers that iterate channel estimation and signal detection on the basis of the expectation-maximization (EM) algorithm have been investigated. In order to make such a MAP receiver more robust against time-variant channels, this paper derives channel estimation with subcarrier removal from the perspective of a message-passing algorithm on factor graphs. The subcarrier removal estimates the channel frequency response by using the detected signals of all subcarriers other than that of a targeted subcarrier. This modification can avoid the repetitive use of incorrectly detected signals for the channel estimation. Computer simulations under fast-fading conditions demonstrate that the EM-based MAP receiver performing the subcarrier removal can improve the packet error rate (PER) drastically.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) is an effective multicarrier modulation technique for high bit-rate transmission over multipath radio channels, because it can transform a frequency-selective fading channel into a set of parallel frequency-flat subchannels. The optimal receiver for the OFDM system is one based on the maximum *a posteriori* (MAP) criterion. Since an ideal receiver of this kind would involve prohibitive computational complexity, a suboptimal receiver is investigated to drastically reduce complexity, but without increasing the packet error rate (PER). The expectation-maximization (EM) algorithm [1], which approximates the MAP estimation in an iterative manner, involves feasible computational complexity, and has been applied to joint signal detection and channel estimation [2]-[5].

In order to track time-varying channels, an EM-based receiver performs the channel estimation using the Kalman filter and employing channel models such as the differential model [2], the random walk model [3], the first-order autoregressive (AR) model [4], or the basis expansion model [5]. However, the channel estimation of this receiver degrades over very fast fading channels. Furthermore, such degradation damages the reliability of the signal detection and vice versa in the iterative process [3]. Especially, higher-order modulation such as 64QAM is vulnerable to the channel estimation error.

For further improvement of the accuracy of the channel estimation, this paper derives channel estimation with subcarrier removal from the perspective of a message-passing algorithm on factor graphs [6], [7]. The subcarrier removal can avoid the repetitive use of incorrectly detected signals for the channel estimation and thus can improve PER of the EM-based receiver.

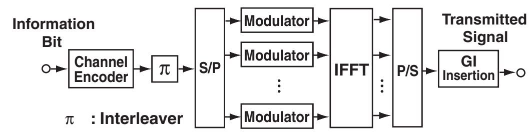

Fig. 1. Structure of an OFDM transmitter

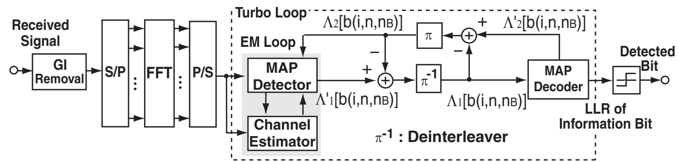

Fig. 2. Structure of an OFDM receiver

II. System Model

A. Transmitter and Receiver Structure

A block diagram of an OFDM transmitter is shown in **Fig. 1**. An information bit sequence is passed through a channel encoder of the convolutional code. The encoded bit sequence is then fed into a 1-symbol-long block interleaver. After the interleaver, the resulting output is transformed into a number of parallel sequences in the serial-parallel converter. These sequences are mapped onto modulation signals when they pass into an inverse fast Fourier transform (IFFT) processor. Finally, the output signals are transformed into OFDM signals with N subcarriers, into which a guard interval (GI) is inserted. Let T_S be the duration of an OFDM symbol. A packet is composed of I OFDM symbols, in which an I_P-symbol-long preamble is followed by $(I - I_P)$ data symbols.

A block diagram of an OFDM receiver is shown in **Fig. 2**. The OFDM signals are transmitted over a multipath Rayleigh-fading channel and are received by the receiver. After removing GI, the resultant signals are fed into an FFT processor. Let us assume that the maximum delay time of the propagation paths is not longer than the duration of GI, and that the timing of the direct propagation path is precisely known. On the assumption that the multipath channel is time-invariant during one symbol, the FFT output $\mathbf{Y}(i)$ at the i-th $(0 \le i < I)$ symbol is expressed in the complex baseband equivalent form as

$$\mathbf{Y}(i) = \mathbf{X}(i)\mathbf{F}\mathbf{h}(i) + \mathbf{N}(i). \tag{1}$$

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

$\mathbf{Y}(i) = [Y_0(i)\ Y_1(i)\ \cdots\ Y_{N-1}(i)]^{\mathrm{T}}$ represents the received signal, where $Y_n(i)$ is the received signal at the n-th ($0 \le n < N$) subcarrier and $^{\mathrm{T}}$ denotes transposition. $\mathbf{h}(i) = [h_0(i)\ h_1(i)\ \cdots\ h_{D-1}(i)]^{\mathrm{T}}$ represents the channel impulse response during the i-th symbol period, where $h_d(i)$ is a complex envelope of the d-th propagation path with delay time of d ($0 \le d < D$) normalized by a sampling period. $\mathbf{N}(i) = [N_0(i)\ N_1(i)\ \cdots\ N_{N-1}(i)]^{\mathrm{T}}$ represents additive white Gaussian noise, where $N_n(i)$ is the noise at the n-th subcarrier with mean zero and variance σ_n^2, and is statistically independent of those at different subcarriers. Thus $\mathbf{N}(i) \sim \mathcal{N}_{\mathbb{C}}[\mathbf{0}_N, \sigma_n^2\mathbf{I}_N]$, where $\mathbf{0}_N$ is the N-by-1 zero vector, $\mathbf{I}_N$ is the N-by-N identity matrix, and $\mathcal{N}_{\mathbb{C}}[\mathbf{m}, \mathbf{V}]$ denotes a complex Gaussian process with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{V}$. $\mathbf{X}(i)$ is an N-by-N signal matrix, and $\mathbf{F}$ is an N-by-D Fourier transform matrix. They are defined as

$$\mathbf{X}(i) = \mathrm{diag}[X_0(i)\ X_1(i)\ \cdots\ X_{N-1}(i)], \tag{2}$$

$$\mathbf{F}^{\mathrm{H}} = [\mathbf{f}_0\ \mathbf{f}_1\ \cdots\ \mathbf{f}_{N-1}], \tag{3}$$

$$\mathbf{f}_n = \left[1\ e^{j2\pi n/N}\ \cdots\ e^{j2\pi n(D-1)/N}\right]^{\mathrm{T}}, \tag{4}$$

where diag[ ] and $^{\mathrm{H}}$ denote a diagonal matrix and Hermitian conjugate, respectively. $X_n(i)$ is a modulation signal at the n-th subcarrier. After FFT, the receiver performs joint iterative processing of (i) channel estimation, (ii) signal detection, and (iii) channel decoding, as will be described in detail later.

B. Channel Model

As the process equation of the Kalman filter, this paper adopts the differential model, which includes the first- and higher-order time differentials of the channel impulse response [2]. Since this model can to some extent realize correlated time-variations, channel estimation using this model offers an improved ability to track fast-fading channels.

Let $h_d^{(u)}(i)$ denote the u-th differential of $h_d(i)$. Approximating $h_d(i)$ up to the $(U-1)$-th differential, the process equation of the d-th propagation path is expressed as

$$\tilde{\mathbf{h}}_d(i+1) = \mathbf{\Phi}_0\tilde{\mathbf{h}}_d(i) + \boldsymbol{\nu}_d(i), \tag{5}$$

$$\mathbf{\Phi}_0 = \begin{bmatrix} 1 & 1/1! & & \cdots & \frac{1}{(U-1)!} \\ 0 & 1 & 1/1! & \cdots & \frac{1}{(U-2)!} \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & & 0 & 1 \end{bmatrix}, \tag{6}$$

$$\tilde{\mathbf{h}}_d(i) = [h_d(i)\ T_S h_d^{(1)}(i)\ \cdots\ T_S^{U-1} h_d^{(U-1)}(i)]^{\mathrm{T}}, \tag{7}$$

$$\boldsymbol{\nu}_d(i) = [\nu_{d,0}(i)\ T_S\nu_{d,1}(i)\ \cdots\ T_S^{U-1}\nu_{d,U-1}(i)]^{\mathrm{T}}, \tag{8}$$

where $\tilde{\mathbf{h}}_d(i)$, $\mathbf{\Phi}_0$, and $\boldsymbol{\nu}_d(i)$ are a U-by-1 state vector of the d-th propagation path, a U-by-U transition matrix, and a U-by-1 process noise vector, respectively. $\nu_{d,u}(i)$ with $1 \le u < U$ is a residual error of the approximation, which contains the U-th and higher-order differentials included in the Taylor series expansions of $h_d^{(u)}(i+1)$ about iT_S.

Furthermore, to express all the path components including their differentials, the process equation using a DU-by-1 extended state vector $\mathbf{h}_{\mathrm{ex}}^{\mathrm{T}}(i) = [\tilde{\mathbf{h}}_0^{\mathrm{T}}(i)\ \tilde{\mathbf{h}}_1^{\mathrm{T}}(i)\ \cdots\ \tilde{\mathbf{h}}_{D-1}^{\mathrm{T}}(i)]$ can be obtained from (5) as

$$\mathbf{h}_{\mathrm{ex}}(i+1) = \mathbf{\Phi}\mathbf{h}_{\mathrm{ex}}(i) + \boldsymbol{\nu}_{\mathrm{ex}}(i), \tag{9}$$

$$\mathbf{\Phi} = \mathbf{I}_D \otimes \mathbf{\Phi}_0 = \begin{bmatrix} \mathbf{\Phi}_0 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{\Phi}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{O} \\ \mathbf{O} & \cdots & \mathbf{O} & \mathbf{\Phi}_0 \end{bmatrix}, \tag{10}$$

$$\boldsymbol{\nu}_{\mathrm{ex}}^{\mathrm{T}}(i) = [\boldsymbol{\nu}_0^{\mathrm{T}}(i)\ \boldsymbol{\nu}_1^{\mathrm{T}}(i)\ \cdots\ \boldsymbol{\nu}_{D-1}^{\mathrm{T}}(i)], \tag{11}$$

where $\otimes$ denotes the Kronecker product, and $\mathbf{\Phi}$ is the DU-by-DU transition matrix. $\boldsymbol{\nu}_{\mathrm{ex}}(i)$ is assumed to follow the multivariate complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}[\mathbf{0}_{DU}, \mathbf{\Sigma}(i)]$ and to be independent of $\mathbf{h}_{\mathrm{ex}}(i)$. The model that uses (9) as the process equation is referred to as the differential model.

Using $\mathbf{h}_{\mathrm{ex}}(i)$, we can rewrite (1) as

$$\mathbf{Y}(i) = \mathbf{X}(i)\mathbf{F}\mathbf{E}\mathbf{h}_{\mathrm{ex}}(i) + \mathbf{N}(i), \tag{12}$$

$$\mathbf{E} = [\mathbf{e}_1\ \mathbf{e}_{U+1}\ \cdots\ \mathbf{e}_{(D-1)U+1}]^{\mathrm{T}}, \tag{13}$$

where $\mathbf{e}_q$ is a DU-by-1 unit vector in which the q-th element is 1 and the others are 0, and $\mathbf{h}(i) = \mathbf{E}\mathbf{h}_{\mathrm{ex}}(i)$.

Although we adopt the differential model, channel estimation proposed in Section V can be also applied to receivers adopting other channel models, such as the random walk model [3] and the first-order AR model [4].

III. MAP Receiver

The MAP receiver, whose structure is shown in Fig. 2, is based on the turbo principle. Specifically, the MAP decoder for channel decoding iteratively exchanges extrinsic information with a soft demodulator that consists of a MAP detector and a channel estimator. The process of the MAP receiver is described below.

During the I_P preamble symbol periods, the initial channel estimation is performed. At the i-th ($i \ge I_P$) data symbol, joint signal detection and channel estimation is iteratively performed via the EM algorithm, which will be detailed in Subsection IV-B. This iteration is referred to as the EM iteration, of which the maximum number is L. After L iterations, the *a posteriori* probability of $X_n(i)$ is calculated using the latest channel estimate. Then, the probability is converted into log likelihood ratios (LLRs) of coded bits, $\Lambda_1'[b(i,n,n_B)]$, where $\{b(i,n,n_B)|1 \le n_B \le N_B\}$ are coded bits to determine $X_n(i)$. n_B and N_B are a bit index and the number of bits of a modulation signal, respectively.

Then, $\Lambda_1'[b(i,n,n_B)]$ of the i-th symbol is fed to the 1-symbol-long block deinterleaver and permuted into $\Lambda_1[b(i,n,n_B)]$. The MAP decoder uses $\Lambda_1[b(i,n,n_B)]$ as *a priori* information for the forward and backward Viterbi decoding in the logarithm domain, and calculates the LLRs of the information bits along the trellis of a convolutional code corresponding to the i-th symbol. Since we cannot know a final state of the trellis which depends on the $(i+1)$-th symbol, the MAP decoder simply assumes equiprobable states as the final state for the Viterbi decoding. For simplicity, an initial state of the trellis which depends on the $(i-1)$-th symbol also can be set equiprobable. The sign of the LLR of each information bit determines the detected bit. The MAP decoder also generates

TABLE I

PROCEDURE OF THE EM ALGORITHM

step1 : Set an initial estimate $\hat{\boldsymbol{\theta}}^0$ and $l = 0$
step2 : Repeat the following operations while $l < L$
E-step : Calculate $Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^l)$
M-step : $\hat{\boldsymbol{\theta}}^{l+1} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^l)$ and set $l \leftarrow l + 1$

the LLRs of the coded bits along the same trellis, which is denoted by $\Lambda_2'[b(i,n,n_B)]$.

Subtracting $\Lambda_1[b(i,n,n_B)]$ from $\Lambda_2'[b(i,n,n_B)]$ and interleaving the resultant output yields *a priori* LLR $\Lambda_2[b(i,n,n_B)]$. Then, $\Lambda_2[b(i,n,n_B)]$ is converted into the *a priori* probability of the modulation signal of the i-th symbol. The EM iteration using the *a priori* probability is performed L times, and then $\Lambda_1'[b(i,n,n_B)]$ is regenerated.

Subtracting $\Lambda_2[b(i,n,n_B)]$ from $\Lambda_1'[b(i,n,n_B)]$ and deinterleaving the resultant output yields $\Lambda_1[b(i,n,n_B)]$. Using $\Lambda_1[b(i,n,n_B)]$, the MAP decoder performs channel decoding again. The operation of the EM iteration and the MAP decoding is referred to as "turbo iteration", in which the iteration index is l_T $(0 \le l_T < L_T)$. After L_T repetitions of the turbo loop, the processing for the $(i+1)$-th symbol is performed using the latest channel estimate of the i-th symbol.

IV. THE EM ALGORITHM

A. Summary of the EM Algorithm

Let $\boldsymbol{\theta}$, $\mathbf{Y}$, and $\mathbf{i}$ represent unknown random parameters, observed data, and hidden variables, respectively. The optimal estimation under the MAP criterion is given by

$$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{Y}). \tag{14}$$

The procedure of the EM algorithm applied to this MAP estimation is summarized in **Table I**, where l $(\le L)$ denotes the iteration index of the EM algorithm and $\hat{\boldsymbol{\theta}}^l$ denotes the l-th estimate. In addition, the function $Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ is given by

$$Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = E_{\mathbf{i}}[\log p(\boldsymbol{\theta}|\mathbf{i}, \mathbf{Y})|\mathbf{Y}, \hat{\boldsymbol{\theta}}], \tag{15}$$

where $E_{\mathbf{i}}[\cdot|\mathbf{Y}, \hat{\boldsymbol{\theta}}]$ is the expectation operation with respect to $\mathbf{i}$, given the occurrence of $\mathbf{Y}$ and $\hat{\boldsymbol{\theta}}$. The EM algorithm ensures that $\log p(\hat{\boldsymbol{\theta}}^l|\mathbf{Y})$ increases monotonically with l [1]. However, convergence to the global optimum is not guaranteed; thus, a bad initial estimate could result in a local maximum.

In this paper, we apply the EM algorithm to the MAP-based channel estimation for OFDM receivers. This application leads to the receiver which iterates soft signal detection and the minimum mean square error (MMSE) channel estimation [2].

B. Joint Signal Detection and Channel Estimation

Let $\mathcal{Y}_0^i = \{\mathbf{Y}(0), \mathbf{Y}(1), \cdots \mathbf{Y}(i)\}$ and $\mathcal{X}_0^i = \{\mathbf{X}(0), \mathbf{X}(1), \cdots \mathbf{X}(i)\}$. In addition, the observed data $\mathbf{Y}$, the unknown random parameter $\boldsymbol{\theta}$, and the hidden variable $\mathbf{i}$ correspond to $\mathcal{Y}_0^i$, $\mathbf{h}_{\text{ex}}(i)$, and $\mathcal{X}_0^i$, respectively. Setting $\mathbf{Y} = \mathcal{Y}_0^i$ leads to a forward recursive form of the channel estimation along the time axis, as will be shown later. With such a setting, the operations of the E-step and M-step are given by

$$\underline{\text{E-step}}: \quad Q[\mathbf{h}_{\text{ex}}(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)]$$
$$= E_{\mathcal{X}_0^i}\left\{\log p[\mathbf{h}_{\text{ex}}(i)|\mathcal{X}_0^i, \mathcal{Y}_0^i]\middle|\mathcal{Y}_0^i, \hat{\mathbf{h}}_{\text{ex}}^l(i)\right\}, \tag{16}$$

$$\underline{\text{M-step}}: \quad \hat{\mathbf{h}}_{\text{ex}}^{l+1}(i) = \arg\max_{\mathbf{h}_{\text{ex}}(i)} Q[\mathbf{h}_{\text{ex}}(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)]. \tag{17}$$

The derivation of the E-step and the M-step is detailed in [2]. The soft decision obtained by the E-step is given by

$$\underline{\text{E-step}}: \quad \langle X_n(i)\rangle = \sum_{m=1}^{M} s_m p[X_n(i) = s_m|Y_n(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)], \tag{18}$$

$$\langle |X_n(i)|^2\rangle = \sum_{m=1}^{M} |s_m|^2 p[X_n(i) = s_m|Y_n(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)], \tag{19}$$

where s_m and M denote the modulation symbol and modulation order. For the sake of simplicity, $\langle\cdot\rangle = E_{X_n(i)}[\cdot|Y_n(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)]$. $p[X_n(i) = s_m|Y_n(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)]$ is calculated as

$$p[X_n(i) = s_m|Y_n(i), \hat{\mathbf{h}}_{\text{ex}}^l(i)]$$
$$= \frac{p[Y_n(i)|\hat{\mathbf{h}}_{\text{ex}}^l(i), X_n(i) = s_m]\ p[X_n(i) = s_m]}{\sum_{\tilde{m}=1}^{M} p[Y_n(i)|\hat{\mathbf{h}}_{\text{ex}}^l(i), X_n(i) = s_{\tilde{m}}]\ p[X_n(i) = s_{\tilde{m}}]}, \tag{20}$$

where $p[Y_n(i)|\hat{\mathbf{h}}_{\text{ex}}^l(i), X_n(i) = s_m]$ is derived as

$$p[Y_n(i)|\hat{\mathbf{h}}_{\text{ex}}^l(i), X_n(i) = s_m] = p[Y_n(i)|\hat{H}_n^l(i), X_n(i) = s_m]$$
$$= \frac{1}{\pi\sigma_n^2}\exp\left[-\frac{|Y_n(i) - s_m\hat{H}_n^l(i)|^2}{\sigma_n^2}\right], \tag{21}$$

$$\hat{H}_n^l(i) = \mathbf{f}_n^{\text{H}}\mathbf{E}\hat{\mathbf{h}}_{\text{ex}}^l(i), \tag{22}$$

using (12), (2), and (3). $\hat{H}_n^l(i)$ is the l-th estimate of the channel frequency response at the n-th subcarrier. Also, $p[X_n(i) = s_m]$ in (20) is the *a priori* probability of $X_n(i)$. The equiprobable symbols are used for $l_T = 0$. When $l_T \ge 1$, the *a priori* probability can be obtained from the LLRs of coded bits, which are provided by the channel decoder. During the preamble $(0 \le i < I_P)$ period, the *a priori* probability is equal to 1 when s_m coincides with the transmitted symbol; otherwise it is 0.

Using the soft decision obtained by the E-step, the operation of the M-step yields the $(l+1)$-th channel estimate as

$$\underline{\text{M-step}}: \quad \hat{\mathbf{h}}_{\text{ex}}^{l+1}(i) = \mathbf{P}(i)\mathbf{v}(i), \tag{23}$$

$$\mathbf{P}^{-1}(i) = \mathbf{P}^{-1}(i|i-1) + \sum_{n=0}^{N-1}\langle|X_n(i)|^2\rangle\mathbf{E}^{\text{H}}\mathbf{f}_n\mathbf{f}_n^{\text{H}}\mathbf{E}, \tag{24}$$

$$\mathbf{v}(i) = \mathbf{P}^{-1}(i|i-1)\hat{\mathbf{h}}_{\text{ex}}(i|i-1) + \sum_{n=0}^{N-1} Y_n(i)\langle X_n^*(i)\rangle\mathbf{E}^{\text{H}}\mathbf{f}_n, \tag{25}$$

where * denotes complex conjugation. $\mathbf{P}(i)$ and $\mathbf{v}(i)$ are a DU-by-DU matrix and a DU-by-1 vector, respectively. $\hat{\mathbf{h}}_{\text{ex}}(i|i-1)$ and $\sigma_n^2\mathbf{P}(i|i-1)$ are the mean and covariance matrix of $\mathbf{h}_{\text{ex}}(i)$, respectively, given the occurrence of $\mathcal{X}_0^{i-1}$ and $\mathcal{Y}_0^{i-1}$. $\hat{\mathbf{h}}_{\text{ex}}(i|i-1)$ and $\mathbf{P}(i|i-1)$ are calculated as

$$\hat{\mathbf{h}}_{\text{ex}}(i|i-1) = \boldsymbol{\Phi}\hat{\mathbf{h}}_{\text{ex}}^L(i-1), \tag{26}$$

$$\mathbf{P}(i|i-1) = \lambda^{-1}\boldsymbol{\Phi}\mathbf{P}^L(i-1)\boldsymbol{\Phi}^{\text{H}}. \tag{27}$$

where $\mathbf{P}^L(i-1) = \mathbf{P}(i-1)$ at the L-th iteration $(l = L)$. λ $(0 < \lambda \le 1)$ is the forgetting factor [8], which leads to an exponentially weighted recursive least-squares (RLS) like algorithm along the time axis. Equation (26) is also used as the initial estimate, that is, $\hat{\mathbf{h}}^{0}_{\mathrm{ex}}(i) = \hat{\mathbf{h}}_{\mathrm{ex}}(i|i-1)$.

Although the inversion of $\mathbf{P}(i)$ must be calculated to obtain $\hat{\mathbf{h}}^{l+1}_{\mathrm{ex}}(i)$ in (23)-(25), the application of the matrix inversion lemma [8] can avoid the matrix inversion. This yields a recursive form of the M-step, which is given by

$$\mathbf{k}(i,n) = \mathbf{P}(i,n-1)\mathbf{E}^{\mathrm{H}}\mathbf{f}_n \times[\langle|X_n(i)|^2\rangle^{-1} + \mathbf{f}^{\mathrm{H}}_n\mathbf{E}\mathbf{P}(i,n-1)\mathbf{E}^{\mathrm{H}}\mathbf{f}_n]^{-1}, \tag{28}$$

$$\hat{\mathbf{h}}_{\mathrm{ex}}(i,n) = \hat{\mathbf{h}}_{\mathrm{ex}}(i,n-1) +\mathbf{k}(i,n)\left[\frac{\langle X^*_n(i)\rangle}{\langle|X_n(i)|^2\rangle}Y_n(i) - \mathbf{f}^{\mathrm{H}}_n\mathbf{E}\hat{\mathbf{h}}_{\mathrm{ex}}(i,n-1)\right], \tag{29}$$

$$\mathbf{P}(i,n) = \mathbf{P}(i,n-1) - \mathbf{k}(i,n)\mathbf{f}^{\mathrm{H}}_n\mathbf{E}\mathbf{P}(i,n-1), \tag{30}$$

where $\mathbf{k}(i,n)$ and $\mathbf{P}(i,n)$ are the DU-by-1 Kalman gain vector and a DU-by-DU matrix, respectively. Setting $\hat{\mathbf{h}}_{\mathrm{ex}}(i,-1) = \hat{\mathbf{h}}_{\mathrm{ex}}(i|i-1)$ and $\mathbf{P}^{-1}(i,-1) = \mathbf{P}^{-1}(i|i-1)$ satisfies $\hat{\mathbf{h}}_{\mathrm{ex}}(i,N-1) = \hat{\mathbf{h}}^{l+1}_{\mathrm{ex}}(i)$ and $\mathbf{P}(i,N-1) = \mathbf{P}(i)$. Equations (28)-(30) can be regarded as an RLS-like algorithm in the frequency domain.

V. Message Passing on Factor Graphs

A. The Sum-Product Algorithm and the EM Algorithm

The factor graph is a bipartite graph which can be used to derive a wide variety of algorithms in signal processing and coding theory [6], [7]. The factor graph decomposes a global function such as a joint probability density function into a product of simpler local functions, and can visualize the relation between the local functions and their arguments called variables. Message passing between these local functions and variables can realize efficient marginalization.

The factor graph of the i-th OFDM symbol is shown in **Fig. 3**, where the index i is omitted for simplicity. This factor graph is used to calculate marginalization of the joint probability of $\mathbf{h}_{\mathrm{ex}}$, $\mathbf{X}$, and $\mathbf{Y}$, of which factorization is expressed as

$$p(\mathbf{h}_{\mathrm{ex}}, \mathbf{X}, \mathbf{Y}) = p(\mathbf{h}_{\mathrm{ex}})\prod_{n=0}^{N-1} p(Y_n|\mathbf{h}_{\mathrm{ex}}, X_n)p(X_n). \tag{31}$$

In Fig. 3, $p(\mathbf{h}_{\mathrm{ex}})$, $p(Y_n|\mathbf{h}_{\mathrm{ex}}, X_n)$, and $p(X_n)$ are local functions symbolized by black squares, $\mathbf{h}_{\mathrm{ex}}$ and X_n are variables symbolized by white circles, and Y_n is treated as a parameter.

The most well-known message passing algorithm on the factor graph is the sum-product algorithm [6]. This algorithm can compute the exact *a posteriori* probability of the variables if the factor graph has no cycle. Solid arrows in Fig. 3 are message flows of the algorithm to calculate the *a posteriori* probability of X_n. Specifically, the downward message to X_n can be regarded as the estimate of the channel frequency response of the n-th subcarrier. This message is calculated by all the incoming messages at $\mathbf{h}_{\mathrm{ex}}$ but the message from X_n.

The EM algorithm also can be interpreted as a message passing algorithm on the factor graph [7]. In contrast to the sum-product algorithm, the downward message of the EM algorithm is calculated by all the incoming messages at $\mathbf{h}_{\mathrm{ex}}$, which also include a dashed arrow in Fig. 3. This can be seen from (23)-(25), in which the information of all

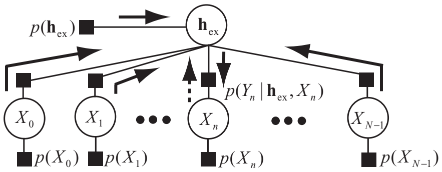

Fig. 3. Message passing on factor graph of OFDM signals

the subcarriers is used. Note that (31) uses $p(\mathbf{h}_{\mathrm{ex}})$ as the *a priori* probability of the impulse response, whereas (23)-(25) uses $p[\mathbf{h}_{\mathrm{ex}}(i)|\mathcal{X}^{i-1}_0, \mathcal{Y}^{i-1}_0]$ instead. Although the EM algorithm makes the message from X_n directly returned to itself, such a returned path intuitively results in the repetitive use of incorrectly detected signals for the channel estimation.

B. Subcarrier Removal

Modification to remove the direct return path from the EM-based channel estimation was introduced in [3]. It first performs channel estimation by the forward and backward RLS algorithm over all the OFDM symbols and then removes the information on the targeted symbol from the channel estimate. Thus, the resultant channel estimate at the i-th OFDM symbol uses information on $\mathcal{Y}^{i-1}_0$ and $\mathcal{Y}^{I}_{i+1}$. Although the method can avoid the repetitive use of incorrectly detected symbols, it additionally needs the backward channel estimation, which leads to doubling of both complexity and latency.

To introduce such removing effect without the backward channel estimation, this paper proposes subcarrier removal. This method estimates the channel frequency response of the n-th subcarrier by using both $\mathcal{Y}^{i-1}_0$ and all the subcarriers of $\mathbf{Y}(i)$ other than $Y_n(i)$. By removing the information of the n-th subcarrier from $\hat{\mathbf{h}}^{l+1}_{\mathrm{ex}}(i)$ which is obtained by (23)-(25), the resultant channel estimate, $\hat{\mathbf{h}}^{l+1}_{\mathrm{ex},\bar{n}}(i)$, is obtained as

$$\begin{aligned}\hat{\mathbf{h}}^{l+1}_{\mathrm{ex},\bar{n}}(i) &= \left[\mathbf{P}^{-1}(i|i-1) + \sum_{n'=0,\ n'\neq n}^{N-1}\langle|X_{n'}(i)|^2\rangle\mathbf{E}^{\mathrm{H}}\mathbf{f}_{n'}\mathbf{f}^{\mathrm{H}}_{n'}\mathbf{E}\right]^{-1} \\ &\times\left[\mathbf{P}^{-1}(i|i-1)\hat{\mathbf{h}}_{\mathrm{ex}}(i|i-1) + \sum_{n'=0,\ n'\neq n}^{N-1} Y_{n'}(i)\langle X^*_{n'}(i)\rangle\mathbf{E}^{\mathrm{H}}\mathbf{f}_{n'}\right] \\ &= [\mathbf{P}^{-1}(i) - \langle|X_n(i)|^2\rangle\mathbf{E}^{\mathrm{H}}\mathbf{f}_n\mathbf{f}^{\mathrm{H}}_n\mathbf{E}]^{-1} \\ &\times[\mathbf{v}(i) - Y_n(i)\langle X^*_n(i)\rangle\mathbf{E}^{\mathrm{H}}\mathbf{f}_n].\end{aligned} \tag{32}$$

To avoid the matrix inversion in (32) for complexity reduction, the matrix inversion lemma [8] is applied. This yields

$$\begin{aligned}&[\mathbf{P}^{-1}(i) - \langle|X_n(i)|^2\rangle\mathbf{E}^{\mathrm{H}}\mathbf{f}_n\mathbf{f}^{\mathrm{H}}_n\mathbf{E}]^{-1} \\ &= \mathbf{P}(i) + \frac{\mathbf{P}(i)\mathbf{E}^{\mathrm{H}}\mathbf{f}_n\mathbf{f}^{\mathrm{H}}_n\mathbf{E}\mathbf{P}(i)}{\langle|X_n(i)|^2\rangle^{-1} - \alpha_n},\end{aligned} \tag{33}$$

where $\alpha_n = \mathbf{f}^{\mathrm{H}}_n\mathbf{E}\mathbf{P}(i)\mathbf{E}^{\mathrm{H}}\mathbf{f}_n$. Thus, $\hat{\mathbf{h}}^{l+1}_{\mathrm{ex},\bar{n}}(i)$ is caluclated as

$$\begin{aligned}\hat{\mathbf{h}}^{l+1}_{\mathrm{ex},\bar{n}}(i) &= \hat{\mathbf{h}}^{l+1}_{\mathrm{ex}}(i) - \frac{\mathbf{P}(i)\mathbf{E}^{\mathrm{H}}\mathbf{f}_n}{\langle|X_n(i)|^2\rangle^{-1} - \alpha_n} \\ &\times\left[\frac{\langle X^*_n(i)\rangle}{\langle|X_n(i)|^2\rangle}Y_n(i) - \mathbf{f}^{\mathrm{H}}_n\mathbf{E}\hat{\mathbf{h}}^{l+1}_{\mathrm{ex}}(i)\right],\end{aligned} \tag{34}$$

using (23) and (32)-(33). Finally, the corresponding channel frequency response, $\hat{H}^{l+1}_{\bar{n}}(i)$, is obtained by (22) and (34) as

TABLE II
SIMULATION PARAMETERS

Modulation	64QAM, 64 subcarreir OFDM
Symbol duration (GI)	80 (16) Δ_t
Packet length	11 (preamble: 1, data: 10)
Channel coding	Convolutional code (R = 1/2, K = 7)
Block interleaver	24×16 (1 Symbol)
Decoding algorithm	Max-Log-MAP
Channel model	16-path Rayleigh with exponential decay
Maximum Doppler frequency	$f_D T_S$ = 0.05

$$\begin{aligned}\hat{H}_{\bar{n}}^{l+1}(i) &= \mathbf{f}_n^{\mathrm{H}}\mathbf{E}\hat{\mathbf{h}}_{\mathrm{ex},\bar{n}}^{l+1}(i)\\ &= \hat{H}_n^{l+1}(i) - \frac{\alpha_n}{\langle|X_n(i)|^2\rangle^{-1} - \alpha_n}\left[\frac{\langle X_n^*(i)\rangle}{\langle|X_n(i)|^2\rangle}Y_n(i) - \hat{H}_n^{l+1}(i)\right].\end{aligned} \tag{35}$$

In summary, the estimate of the channel frequency response with subcarrier removal can be calculated as follows: The receiver performs recursive processes of (28)-(30) to obtain $\hat{\mathbf{h}}_{\mathrm{ex}}^{l+1}(i)$, and next converts it to $\hat{H}_n^{l+1}(i)$. Then, the receiver calculates $\hat{H}_{\bar{n}}^{l+1}(i)$ by (35) for the subcarrier removal. This procedure can calculate $\hat{H}_{\bar{n}}^{l+1}(i)$ efficiently compared to simple calculation of (32) for all the subcarriers.

VI. COMPUTER SIMULATION

A. Simulation Conditions

Computer simulations were conducted in order to verify the performance of the proposed MAP receiver performing the subcarrier removal. **Table II** shows the simulation parameters. 64QAM was assumed as the modulation scheme. A convolutional code with a coding rate $R = 1/2$ and constraint length $K = 7$ was used as the channel coding. As a process equation of the channel estimation, the differential model with the maximum order of the differentials being 1 ($U = 2$) was used. From the results in [2], the forgetting factor λ and the number of EM iterations L were set to 0.3 and 3, respectively. The normalized maximum Doppler frequency was set to $f_D T_S = 5.0 \times 10^{-2}$. As a lower bound, the performance of the receiver under the ideal channel estimation condition was also evaluated.

B. Average PER Performance

Fig. 4 shows the average PER performances of the channel estimation with the subcarrier removal and that with the original EM algorithm. L_T was set to 2 and 10. Due to the higher-order modulation under the fast fading, both the channel estimation with the subcarrier removal and that by the original EM algorithm suffer from an error floor even with $L_T = 10$. However, it can be seen that the modification by the subcarrier removal can improve the PER. The subcarrier removal reduces the error floor from 9.9×10^{-3} to 3.9×10^{-3} with $L_T = 10$. This improvement is due to avoiding the repetitive use of incorrect modulation signals for the channel estimation. Convergence speed of the turbo processing can be seen in **Fig. 5**, which shows the average PER performance versus the turbo iterations with average $E_b/N_0 = 16$ dB and 30 dB. The graph also shows that subcarrier removal can reduce the PER floor more with higher E_b/N_0.

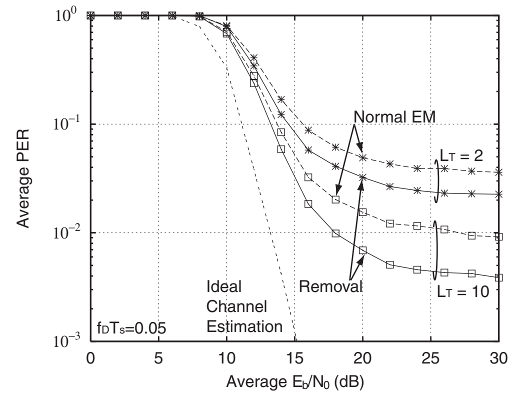

Fig. 4. PER performance versus E_b/N_0

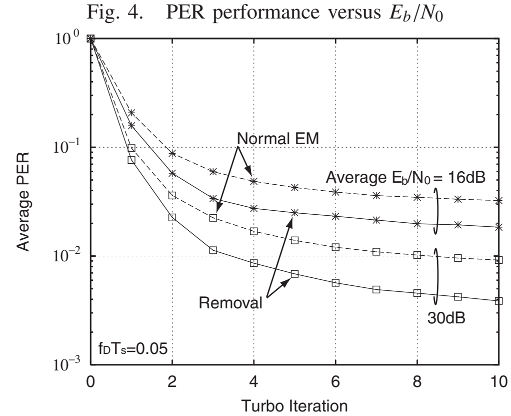

Fig. 5. PER performance versus L_T

VII. CONCLUSIONS

This paper proposed the channel estimation with subcarrier removal for the OFDM receiver. The method estimates the channel frequency response by using all subcarriers other than a targeted subcarrier. Computer simulations with $f_D T_S = 5.0 \times 10^{-2}$ demonstrated that the subcarrier removal with $L_T = 10$ can reduce the error floor to 3.9×10^{-3} with $E_b/N_0 = 30$ dB.

REFERENCES

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. Series B*, vol. 39, no. 1, pp. 1-38, 1977.

[2] K. Muraoka, K. Fukawa, H. Suzuki, and S. Suyama, "Joint signal detection and channel estimation using differential models via EM algorithm for OFDM mobile communications," *IEICE Trans. Commun.*, vol. E94-B, no. 2, pp. 533-545, Feb. 2011.

[3] T. Kashima, K. Fukawa, and H. Suzuki, "Adaptive MAP receiver via the EM algorithm and message passings for MIMO-OFDM mobile communications," *IEEE J. Select. Areas Commun.*, vol. 24, No. 3, pp. 437-447, Mar. 2006.

[4] T. Y. Al-Naffouri, "An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3924-3930, July 2007.

[5] S.-J. Hwang and P. Schniter, "EM-based soft noncoherent equalization of doubly selective channels using tree search and basis expansion," in *Proc. 10th IEEE Signal Processing Advances in Wireless Commun.*, pp. 6-10, June 2009.

[6] 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.

[7] J. Dauwels, S. Korl, and H.-A. Loeliger, "Expectation maximization as message passing," in *Proc. IEEE Inf. Theory*, pp.583-586, Sept. 2005.

[8] S. Haykin, *Adaptive Filter Theory, 3rd ed.*, Prentice-Hall, 1996.