

RESEARCH

Open Access

Channel estimation and optimal training with the LMMSE criterion for OFDM-based two-way relay networks

Minh Tam Tran¹, Jin Soo Wang¹, Ickho Song² and Yun Hee Kim^{1*}

Abstract

In this paper, we consider the linear minimum mean square error (LMMSE) estimation of channel for the two-way relay network employing analog network coding and orthogonal frequency division multiplexing. The channel responses required for self-interference cancelation and coherent detection are estimated in the time domain with the preamble. We derive the optimal training condition for minimizing the mean square error (MSE) of the LMMSE estimator and obtain the corresponding MSE. It is observed that the LMMSE estimator has the optimal training condition equivalent to that of the least square estimator and that the former is less sensitive to the training sequences than the latter. We also propose new training sequences not only satisfying the optimal training condition but also providing the minimum peak-to-average power ratio.

Keywords: Channel estimation; Minimum mean square error; Orthogonal frequency division multiplexing; Training sequences; Two-way relay

1 Introduction

Relay communication has been intensively studied so far due to its capability of improving coverage and reliability at low cost [1-3]. However, the additional transmission phase required for relaying incurs spectral inefficiency in practice. By adopting analog network coding (ANC) which spends two transmission phases in data exchange [4,5], the inefficiency has recently been partially mitigated in two-way relaying (TWR) networks: In the scheme, two communicating source nodes transmit signals simultaneously in the first phase called the multiple access (MAC) phase, where the two signals autonomously form the ANC in the air. In the second phase called the broadcast (BC) phase, the signals received at the relay node are amplified and broadcasted so that the source nodes equipped with self-interference cancelation (SIC) can retrieve the information which originated from their counterparts. Although various TWR networks based on the ANC have been shown to improve the performance theoretically,

they assume perfect channel state information in performing SIC and coherent detection [6-8].

Recently, the problem of channel estimation has been addressed in the ANC-based TWR networks when the channel reciprocity holds between the MAC and BC phases [9-13]. These studies deal with the estimation of the composite channels of the MAC and BC phases at the communicating nodes. Specifically, the maximum likelihood and linear maximum signal-to-noise ratio (SNR) estimators have been developed in flat fading channels in [9]. For orthogonal frequency division multiplexing (OFDM) systems in frequency-selective fading channels, the least square (LS) estimator and optimal training sequences to minimize the mean square error (MSE) have been derived in [10,11], which have subsequently been extended for the joint estimation of carrier frequency offset and channel responses in [13]. The LS estimators, derived for block-based (or equivalently, preamble-based) signals in the time domain [10] and in the frequency domain [11], are shown to require the same optimal training condition and to provide the same MSE. Some training sequences satisfying the optimal training condition with low peak-to-average power ratio (PAPR), preferred for the low-cost and low-power design of OFDM systems

*Correspondence: yheekim@khu.ac.kr

¹Department of Electronics and Radio Engineering, Kyung Hee University, Yongin Gyeonggi-do 446-701, Korea

Full list of author information is available at the end of the article

[14], have also been suggested in [11]. It should be noted that the channel estimation schemes considered for the OFDM-based TWR networks in [10,11,13] are all based on the LS criterion.

Without channel reciprocity, two channel estimation methods, the independent estimation of the MAC and BC channels at the relay and communication nodes [15] and the combined estimation of the MAC and BC channels at the communication nodes [16], have been studied for OFDM-based TWR networks using ANC. It should be mentioned that the former requires the feedback of channel estimates from the relay for the communication nodes to perform SIC and coherent detection. To avoid such a feedback, the latter estimates the composite channels of the MAC and BC phases as in the channel estimators developed for the OFDM-based TWR networks with channel reciprocity [10,11,13]. The channel estimator in [16], which is equivalent to the LS estimator [10,11,13] applied to the case with no channel reciprocity, reduces the complexity of the LS estimator by adopting the method developed for point-to-point OFDM systems [17].

In this paper, the linear minimum mean square error (LMMSE) estimator is adopted for the improvement of channel estimation of OFDM-based TWR networks employing ANC. As in [4-13], we assume the channel reciprocity, which improves the performance of the TWR networks as observed in [18]. We would nonetheless like to mention that the results herein can be easily extended to the networks without channel reciprocity. For the LMMSE estimator with the time-domain signals, we derive the optimal training condition minimizing the MSE and analyze the effect of multipath intensity profile (MIP) on the MSE. It is verified that the optimal training condition for the LMMSE estimation is equivalent to that for the LS estimation. In addition, we propose new sets of training sequences satisfying the optimal training condition and, at the same time, exhibiting the lowest PAPR.

The rest of this paper is organized as follows: Section 2 introduces the system model of the OFDM-based TWR network employing the ANC, for which the channel

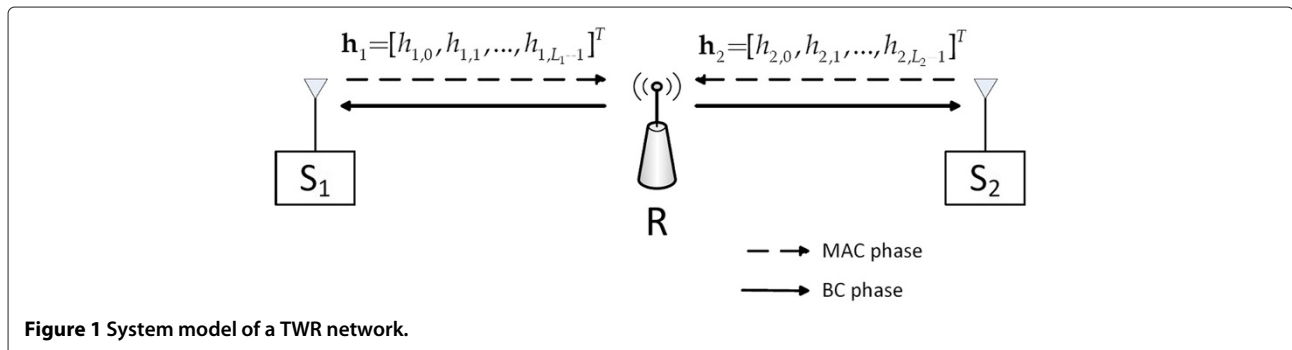
estimation and optimal training condition based on the LMMSE criterion are derived in Section 3. The MSE under the optimal training condition is analyzed in Section 4, followed by discussions on several training sequences satisfying the optimal training condition in Section 5. The performance characteristics of the channel estimators with optimal training sequences are evaluated in Section 6. Section 7 concludes this paper.

2 System model

Consider the TWR network illustrated in Figure 1, where the two source nodes S_1 and S_2 exchange information via the relay node R , and each of the three nodes is equipped with a single antenna: We assume a direct channel is not available between S_1 and S_2 . The information exchange is accomplished over two transmission phases using ANC. In the first phase called the MAC phase, the two source nodes S_1 and S_2 send packets to R simultaneously so that the ANC of two packets is formed autonomously in the received signal at the relay node. In the second phase called the BC phase, the relay R amplifies and broadcasts the signal received in the first phase to S_1 and S_2 . Once the relay signal is received, the nodes S_1 and S_2 perform channel estimation with a preamble and then detect the data symbols from their counter part after SIC.

The channel between S_i and R is assumed to be reciprocal, frequency-selective fading, and quasi-static over two transmission phases for $i = 1, 2$. The discrete-time channel impulse response (CIR) between S_i and R can then be modeled as $\mathbf{h}_i = [h_{i,0} \ h_{i,1} \ \dots \ h_{i,L_i-1}]^T$, where L_i is the number of resolvable multipath taps and $h_{i,l}$ is the complex fading gain of the l th path. Assuming Rayleigh fading, we have $h_{i,l} \sim \mathcal{CN}(0, \sigma_{h_{i,l}}^2)$, where \sim signifies 'distributed as' and $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the circularly symmetric complex Gaussian (CSCG) distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The MIP of the CIR between S_i and R is then described by $\{\sigma_{h_{i,l}}^2\}_{l=0}^{L_i-1}$. The average transmit powers of S_1 , S_2 , and R are denoted by P_1 , P_2 , and P_r , respectively.

We adopt the OFDM signal model with N subcarriers for the block-based channel estimation as depicted,



for example, in [10]. For transmission in the MAC phase, source node S_i generates a preamble followed by a number of vectors of OFDM data symbols: Let $\mathbf{x}_i = [x_{i,0} \ x_{i,1} \ \cdots \ x_{i,N-1}]^T$ and $\mathbf{x}_{i,d} = [x_{i,d,0} \ x_{i,d,1} \ \cdots \ x_{i,d,N-1}]^T$ denote the time-domain representations of the preamble constructed by a training sequence and of the OFDM data symbol vector, respectively, of S_i . The training sequences \mathbf{x}_1 and \mathbf{x}_2 are both known to S_1 and S_2 and are subject to the transmit power constraint $\|\mathbf{x}_1\|^2 = NP_1$ and $\|\mathbf{x}_2\|^2 = NP_2$. Here, the OFDM data symbol vector $\mathbf{x}_{i,d}$ is generated as $\mathbf{x}_{i,d} = \mathbf{W}^H X_{i,d}$, where $X_{i,d} = [X_{i,d,0} \ X_{i,d,1} \ \cdots \ X_{i,d,N-1}]^T$ is the frequency-domain modulation symbol vector with $E\{|X_{i,d,k}|^2\} = P_i$, \mathbf{W} is the unitary discrete Fourier transform (DFT) matrix with the (m,n) th entry $[W]_{m,n} = \frac{1}{\sqrt{N}} \exp(-j2\pi \frac{mn}{N})$, and $(\cdot)^H$ denotes the conjugate transpose. To avoid inter-symbol interference (ISI) in frequency-selective fading channels, a cyclic prefix (CP) of length G is appended, producing $\tilde{\mathbf{x}}_i = [x_{i,N-G}, x_{i,N-G+1}, \dots, x_{i,N-1}, x_{i,0}, x_{i,1}, \dots, x_{i,N-1}]^T$ from the preamble \mathbf{x}_i and $\tilde{\mathbf{x}}_{i,d} = [x_{i,d,N-G}, x_{i,d,N-G+1}, \dots, x_{i,d,N-1}, x_{i,d,0}, x_{i,d,1}, \dots, x_{i,d,N-1}]^T$ from the OFDM data symbol vector $\mathbf{x}_{i,d}$. Normally, we set $G \geq \max(L_1, L_s)$ so that the ISI can be eliminated completely. For the ANC, the two source nodes S_1 and S_2 simultaneously transmit their packets of the CP-added preamble and OFDM data symbol vectors.

The signals received at the relay R in the MAC phase of the ANC protocol can be described, after the removal of CP, as

$$\mathbf{y}_r = \mathbf{\Omega}_1 \mathbf{x}_1 + \mathbf{\Omega}_2 \mathbf{x}_2 + \mathbf{n}_r \quad (1)$$

and

$$\mathbf{y}_{r,d} = \mathbf{\Omega}_1 \mathbf{x}_{1,d} + \mathbf{\Omega}_2 \mathbf{x}_{2,d} + \mathbf{n}_{r,d}, \quad (2)$$

where $\mathbf{\Omega}_i$ is an $N \times N$ column-wise circulant matrix having $\tilde{\mathbf{h}}_i = [\mathbf{h}_i^T \ \mathbf{0}_{1 \times (N-L_i)}]^T$ as its first column, and $\mathbf{n}_r \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_n^2 \mathbf{I}_N)$ and $\mathbf{n}_{r,d} \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_n^2 \mathbf{I}_N)$ are the additive noise vectors at the relay. Here, $\mathbf{0}_{a \times b}$ denotes the all-zero matrix of size $a \times b$ and \mathbf{I}_a is the $a \times a$ identity matrix. The relay constructs $\tilde{\mathbf{y}}_r$ and $\tilde{\mathbf{y}}_{r,d}$ by adding a CP of length G to \mathbf{y}_r and $\mathbf{y}_{r,d}$, respectively, and then broadcasts the amplified signals $\tilde{\mathbf{x}}_r = \alpha \tilde{\mathbf{y}}_r$ and $\tilde{\mathbf{x}}_{r,d} = \alpha \tilde{\mathbf{y}}_{r,d}$ in the BC phase. Here, $\alpha = \sqrt{\frac{P_r}{\beta_1 P_1 + \beta_2 P_2 + \sigma_n^2}}$ is the scaling factor to make the average relay power equal to P_r with $\beta_i = \sum_{l=0}^{L_i-1} \sigma_{h_{i,l}}^2$ for $i = 1, 2$.

Let us now consider the signal processing at node S_i in the BC phase of the ANC protocol. After the removal of

CP, the received signal at S_i can be expressed in the time domain as

$$\mathbf{y}_i = \alpha \mathbf{\Omega}_i \mathbf{\Omega}_i \mathbf{x}_i + \alpha \mathbf{\Omega}_i \mathbf{\Omega}_j \mathbf{x}_j + \alpha \mathbf{\Omega}_i \mathbf{n}_r + \mathbf{n}_i \quad (3)$$

and

$$\mathbf{y}_{i,d} = \alpha \mathbf{\Omega}_i \mathbf{\Omega}_i \mathbf{x}_{i,d} + \alpha \mathbf{\Omega}_i \mathbf{\Omega}_j \mathbf{x}_{j,d} + \alpha \mathbf{\Omega}_i \mathbf{n}_{r,d} + \mathbf{n}_{i,d} \quad (4)$$

for $(i,j) \in \{(1,2), (2,1)\}$, where $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_n^2 \mathbf{I}_N)$ and $\mathbf{n}_{i,d} \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_n^2 \mathbf{I}_N)$ are the additive noise vectors at S_i . Note that $\mathbf{\Omega}_i \mathbf{\Omega}_m$ is a $N \times N$ column-wise circulant matrix having

$$\tilde{\mathbf{t}}_{im} = [\mathbf{t}_{im}^T \ \mathbf{0}_{1 \times (N-\tilde{L}_{im})}]^T = \tilde{\mathbf{h}}_i \circledast \tilde{\mathbf{h}}_m \quad (5)$$

as its first column for $m = i, j$, where \circledast denotes the circular convolution^a, and the vector \mathbf{t}_{im} of length $\tilde{L}_{im} = L_i + L_m - 1$ denotes the non-zero portion of the composite channel represented by the circular convolution of the zero-padded CIRs $\tilde{\mathbf{h}}_i$ and $\tilde{\mathbf{h}}_m$. In (5), the l th element of \mathbf{t}_{im} is given by

$$t_{im,l} = \sum_{l'=\max(0,l-L_i+1)}^{\min(l,L_m-1)} h_{i,l-l'} h_{m,l'} \quad (6)$$

for $l = 0, 1, \dots, \tilde{L}_{im} - 1$ and $m = i, j$.

To detect the modulation symbol vector $X_{j,d}$ transmitted from S_j , the DFT is performed on $\mathbf{y}_{i,d}$ as $\mathbf{Y}_{i,d} = \mathbf{W} \mathbf{y}_{i,d}$, resulting in

$$\mathbf{Y}_{i,d} = \alpha \sqrt{N} \text{diag}(\mathbf{W}_{\tilde{L}_{ii}} \mathbf{t}_{ii}) \mathbf{X}_{i,d} + \alpha \sqrt{N} \text{diag}(\mathbf{W}_{\tilde{L}_{ij}} \mathbf{t}_{ij}) \mathbf{X}_{j,d} + \alpha \mathbf{W} \mathbf{\Omega}_i \mathbf{n}_{r,d} + \mathbf{W} \mathbf{n}_{i,d}, \quad (7)$$

where $\text{diag}(\mathbf{a})$ denotes the diagonal matrix with the vector \mathbf{a} on the diagonal, \mathbf{W}_L is the $N \times L$ submatrix of \mathbf{W} constructed by the first L columns, and we have used $\mathbf{W} \mathbf{\Omega}_i \mathbf{\Omega}_m \mathbf{W}^H = \sqrt{N} \text{diag}(\mathbf{W} \tilde{\mathbf{t}}_{im}) = \sqrt{N} \text{diag}(\mathbf{W}_{\tilde{L}_{im}} \mathbf{t}_{im})$ from the property of a circulant matrix. In (7), the first term on the right-hand side (RHS) is the self-interference (SI) containing S_i 's own transmitted signal, and the second term contains the desired signal $\mathbf{X}_{j,d}$. Hence, S_i first removes the SI from $\mathbf{Y}_{i,d}$ as $\tilde{\mathbf{Y}}_{i,d} = \mathbf{Y}_{i,d} - \alpha \sqrt{N} \text{diag}(\mathbf{W}_{\tilde{L}_{ii}} \mathbf{t}_{ii}) \mathbf{X}_{i,d}$ with channel state information \mathbf{t}_{ii} and then detects the modulation symbol vector $\mathbf{X}_{j,d}$ from $\tilde{\mathbf{Y}}_{i,d}$ with channel state information \mathbf{t}_{ij} . To this end, node S_i of the ANC-based TWR network is required to estimate \mathbf{t}_{ii} and \mathbf{t}_{ij} from the received signal (3) by exploiting the knowledge on the training sequences \mathbf{x}_i and \mathbf{x}_j .

3 Channel estimation and optimal training condition with the LMMSE criterion

Taking into account that the matrices $\Omega_i \Omega_i$ and $\Omega_i \Omega_j$ are circulant, we can rewrite (3) as

$$\begin{aligned} y_i &= \alpha [\Psi_{ii} \Psi_{ij}] \begin{bmatrix} t_{ii} \\ t_{ij} \end{bmatrix} + \alpha \Omega_i n_r + n_i \\ &= \alpha \Psi_i t_i + \tilde{n}_i, \end{aligned} \quad (8)$$

where $\Psi_i = [\Psi_{ii} \Psi_{ij}]$ with Ψ_{im} , the $N \times \tilde{L}_{im}$ column-wise circulant matrix having x_m as its first column; $t_i = [t_{ii}^T \ t_{ij}^T]^T$ denotes the total composite channel to be estimated; and $\tilde{n}_i = \alpha \Omega_i n_r + n_i$. It should be mentioned that the first part t_{ii} and the second part t_{ij} of the total composite channel t_i are essential elements in performing the SIC and coherent demodulation, respectively.

Let us first derive the LMMSE estimate

$$\hat{t}_i = [\hat{t}_{ii}^T \ \hat{t}_{ij}^T]^T = K y_i \quad (9)$$

of the total composite channel t_i in such a way that the MSE

$$\text{MSE}(\hat{t}_i) = E \left\{ \|t_i - \hat{t}_i\|^2 \right\} \quad (10)$$

is minimized. The linear transformation matrix K can be obtained in a straightforward manner from the orthogonality principle [19]

$$\frac{\partial}{\partial K} \text{MSE}(\hat{t}_i) = E \left\{ (t_i - \hat{t}_i) y_i^H \right\} = 0, \quad (11)$$

which leads to $K = R_{t_i y_i} R_{y_i}^{-1}$, or equivalently,

$$\hat{t}_i = R_{t_i y_i} R_{y_i}^{-1} y_i. \quad (12)$$

In (12), $R_{t_i y_i} = E \{ t_i y_i^H \} = \alpha R_{t_i} \Psi_i^H$ is the cross correlation between t_i and y_i , and $R_{y_i} = E \{ y_i y_i^H \} = \alpha^2 \Psi_i R_{t_i} \Psi_i^H + R_{\tilde{n}_i}$ is the autocorrelation of y_i with $R_{t_i} = E \{ t_i t_i^H \}$ denoting the autocorrelation of t_i .

Now, noting that the autocorrelation $R_{\tilde{n}_i}$ of noise \tilde{n}_i can be expressed as

$$R_{\tilde{n}_i} = E \{ \tilde{n}_i \tilde{n}_i^H \} = \sigma_n^2 (\alpha^2 \beta_i + 1) I_N \quad (13)$$

since $E \{ \Omega_i \Omega_i^H \} = \beta_i I_N$, the LMMSE estimator \hat{t}_i in (12) can be rewritten equivalently as

$$\hat{t}_i = \alpha R_{t_i} \Psi_i^H \{ \alpha^2 \Psi_i R_{t_i} \Psi_i^H + \sigma_n^2 (\alpha^2 \beta_i + 1) I_N \}^{-1} y_i \quad (14)$$

and produces the MSE

$$\text{MSE}(\hat{t}_i) = \text{tr} \left\{ \left(R_{t_i}^{-1} + \theta_i \Psi_i^H \Psi_i \right)^{-1} \right\}, \quad (15)$$

where $\text{tr}\{\cdot\}$ denotes the trace of a matrix and

$$\begin{aligned} \theta_i &= \frac{\alpha^2}{\sigma_n^2 (\alpha^2 \beta_i + 1)} \\ &= \frac{P_r}{\sigma_n^2 (\beta_1 P_1 + \beta_2 P_2 + \beta_i P_r + \sigma_n^2)}. \end{aligned} \quad (16)$$

It should be noted that the MSE (15) can be decomposed into $\text{MSE}(\hat{t}_i) = \text{MSE}(\hat{t}_{i1}) + \text{MSE}(\hat{t}_{i2})$, where

$$\text{MSE}(\hat{t}_{im}) = E \left\{ \|t_{im} - \hat{t}_{im}\|^2 \right\} \quad (17)$$

will be called the partial MSE for $m = 1, 2$ in the sequel.

Let us next obtain an explicit expression of R_{t_i} for use in (14) and (15). First, since h_1 and h_2 are independent of each other, it is immediate to have $R_{t_i} = \begin{bmatrix} R_{t_{ii}} & 0 \\ 0 & R_{t_{ij}} \end{bmatrix}$, where $R_{t_{im}} = E \{ t_{im} t_{im}^H \}$. Next, the (l_1, l_2) th element $[R_{t_{im}}]_{l_1, l_2} = E \{ t_{im, l_1} t_{im, l_2}^* \}$ of $R_{t_{im}}$ can be expressed as

$$\begin{aligned} [R_{t_{im}}]_{l_1, l_2} &= \sum_{l'_1 = \max(0, l_1 - L_i + 1)}^{\min(l_1, L_i - 1)} \sum_{l'_2 = \max(0, l_2 - L_i + 1)}^{\min(l_2, L_i - 1)} \\ &\quad \times E \left\{ h_{i, l_1 - l'_1} h_{m, l'_1} h_{i, l_2 - l'_2}^* h_{m, l'_2}^* \right\} \end{aligned} \quad (18)$$

from (6). Here, since $\{h_{i, l}\}_{l=0}^{L_i-1}$ are independent zero-mean CSCG random variables, we have $E \{ h_{i, l}^2 \} = E \{ (h_{i, l}^*)^2 \} = 0$ for all l , $E \{ h_{i, l} h_{i, l'} \} = E \{ h_{i, l} \} E \{ h_{i, l'} \} = 0$ for $l \neq l'$, and $E \{ h_{1, l} h_{2, l'} \} = E \{ h_{1, l} \} E \{ h_{2, l'} \} = 0$ for all l and l' . Therefore, the summand $E \{ h_{i, l_1 - l'_1} h_{m, l'_1} h_{i, l_2 - l'_2}^* h_{m, l'_2}^* \}$ on the RHS of (18) has a non-zero value only when either $l_1 - l'_1 = l_2 - l'_2$ and $l'_1 = l'_2$ or $l_1 - l'_1 = l'_2$ and $l'_1 = l_2 - l'_2$ if $i = m$, and only when $l_1 - l'_1 = l_2 - l'_2$ and $l'_1 = l'_2$ if $i \neq m$. In other words, we have

$$[R_{t_{im}}]_{l_1, l_2} = 0 \quad (19)$$

for $l_1 \neq l_2$, and

$$[R_{t_{ii}}]_{l, l} \triangleq \zeta_{ii, l} = \begin{cases} 2 \sum_{l' = \max(0, l - L_i + 1)}^{\min(l, L_i - 1)} E \{ |h_{i, l'}|^2 \} E \{ |h_{i, l - l'}|^2 \}, & \text{for } l \text{ odd,} \\ 2 \sum_{\substack{l' = \max(0, l - L_i + 1) \\ l' \neq \frac{l}{2}}}^{\min(l, L_i - 1)} E \{ |h_{i, l'}|^2 \} E \{ |h_{i, l - l'}|^2 \} + E \left\{ \left| h_{i, \frac{l}{2}} \right|^4 \right\}, & \text{for } l \text{ even} \end{cases} \quad (20)$$

and

$$[\mathbf{R}_{t_{ij}}]_{l,l} \triangleq \zeta_{ij,l} = \sum_{l'=\max(0,l-L_i+1)}^{\min(l,L_i-1)} E\{|h_{i,l'}|^2 |h_{i,l-l'}|^2\} \quad (21)$$

for $l = 0, 1, \dots, L_i - 1$. Noting that $E\{|h_{i,l}|^4\} = 2\sigma_{h_{i,l}}^2$ and $E\{|h_{i,l}|^2\} = \sigma_{h_{i,l}}^2$, we can combine (20) and (21) to obtain

$$\zeta_{im,l} = (1 + \delta_{i,m}) \sum_{l'=\max(0,l-L_i+1)}^{\min(l,L_m-1)} \sigma_{h_{i,l'}}^2 \sigma_{h_{m,l-l'}}^2, \quad (22)$$

where $\delta_{i,m} = 1$ for $i = m$, and 0 for $i \neq m$ is the Kronecker delta. In short, we finally have

$$\mathbf{R}_{t_{im}} = \text{diag}(\boldsymbol{\zeta}_{im}), \quad (23)$$

where $\boldsymbol{\zeta}_{im} = [\zeta_{im,0} \ \zeta_{im,1} \ \dots \ \zeta_{im,\tilde{L}_{im}-1}]^T$.

Theorem 1. *The training sequences \mathbf{x}_1 and \mathbf{x}_2 , minimizing $\text{MSE}(\hat{\mathbf{t}}_1)$ and $\text{MSE}(\hat{\mathbf{t}}_2)$ subject to $\|\mathbf{x}_1\|^2 = NP_1$ and $\|\mathbf{x}_2\|^2 = NP_2$, should satisfy the zero autocorrelation condition except at the origin*

$$\sum_{n=0}^{N-1} x_{i,n} x_{i,\text{mod}(n+l,N)}^* = \begin{cases} NP_i, & l = 0, \\ 0, & |l| = 1, 2, \dots, \tilde{L}_i - 1 \end{cases} \quad (24)$$

for $i = 1, 2$ and the zero cross-correlation condition

$$\sum_{n=0}^{N-1} x_{1,n} x_{2,\text{mod}(n+l,N)}^* = 0, \quad |l| = 0, 1, \dots, \tilde{L}_{\max} - 1, \quad (25)$$

where $\text{mod}(x, y)$ is the remainder when x is divided by y , $\tilde{L}_i = \max(\tilde{L}_{i1}, \tilde{L}_{i2})$, and $\tilde{L}_{\max} = \max(\tilde{L}_{11}, \tilde{L}_{12}, \tilde{L}_{22}) = 2 \max(L_1, L_2) - 1$ with $\tilde{L}_{im} = L_i + L_m - 1$ for $m = 1, 2$.

Proof. See Appendix 1. \square

4 The MSE of the LMMSE estimator with optimal training sequences

With training sequences satisfying (24) and (25), the MSE (15) of the LMMSE estimator at S_i has the minimum (optimal) value

$$\text{MSE}_o(\hat{\mathbf{t}}_i) = \text{MSE}_o(\hat{\mathbf{t}}_{i1}) + \text{MSE}_o(\hat{\mathbf{t}}_{i2}), \quad (26)$$

where the optimal partial MSE

$$\begin{aligned} \text{MSE}_o(\hat{\mathbf{t}}_{im}) &= \text{tr} \left\{ \left(\mathbf{R}_{t_{im}}^{-1} + \theta_i NP_m \mathbf{I}_{\tilde{L}_{im}} \right)^{-1} \right\} \\ &= \sum_{l=0}^{\tilde{L}_{im}-1} \frac{\zeta_{im,l}}{1 + N\theta_i P_m \zeta_{im,l}} \end{aligned} \quad (27)$$

is the minimum value of the partial MSE (17) incurred in estimating \mathbf{t}_{im} for $m = 1, 2$. It is obvious from (26) and (27) that the optimal MSE of the LMMSE estimator achieved with optimal training sequences would in general depend on the DFT size, powers of signal and noise, and MIPs $\{\sigma_{h_{i,l}}^2\}_{l=0}^{L_s-1}$ which determine \tilde{L}_{im} , $\zeta_{im,l}$, and θ_i .

Now, from (27), we have

$$\begin{aligned} \text{MSE}_o(\hat{\mathbf{t}}_{im}) &\approx \frac{1}{\theta_i} \frac{\tilde{L}_{im}}{NP_m} \\ &= \sigma_n^2 \left(\beta_i + \frac{\beta_1 P_1 + \beta_2 P_2 + \sigma_n^2}{P_r} \right) \frac{\tilde{L}_{im}}{NP_m} \end{aligned} \quad (28)$$

if $\theta_i P_m N \gg 1$ for $m = i, j$ (or equivalently, if $\frac{P_1}{\sigma_n^2} \gg 1$, $\frac{P_2}{\sigma_n^2} \gg 1$, and $\frac{P_r}{\sigma_n^2} \gg 1$), and

$$\text{MSE}_o(\hat{\mathbf{t}}_{im}) \approx \sum_{l=0}^{\tilde{L}_{im}-1} \zeta_{im,l} = (1 + \delta_{i,m}) \beta_i \beta_m \quad (29)$$

if $\frac{P_1}{\sigma_n^2} \ll 1$, $\frac{P_2}{\sigma_n^2} \ll 1$, and $\frac{P_r}{\sigma_n^2} \ll 1$. It should be noted, as described in Appendix 2, that the optimal MSE (26) with the high SNR approximation (28) of the partial MSE for the LMMSE estimation is the same as the optimal MSE of the LS estimation [9]. It is also interesting to note, from (28) and (29), that the numbers L_1 and L_2 of resolvable multipaths relative to the DFT size N are the key factors determining the optimal MSE in the high SNR region, while the average channel powers β_1 and β_2 are the key factors determining the optimal MSE in the low SNR region.

5 New sets of optimal training sequences

In Appendix 2, we have provided a derivation of the LS estimator and its optimal training condition to minimize the MSE, where we have also shown the equivalence of the optimal training conditions for the LMMSE and LS estimators. Interestingly, although the LMMSE estimator is a function of the SNR unlike the LS estimator, the optimal training condition of the LMMSE estimator, like the LS estimator, does not depend on the SNR.

The equivalence observation on the training conditions allows us to conveniently reuse, in the LMMSE estimation, the training sequences proposed originally for the LS estimation. For example, the sequences proposed in

[10] are given with the frequency-domain representation $X_i = \mathbf{W}\mathbf{x}_i = [X_{i,0} \ X_{i,1} \ \dots \ X_{i,N-1}]$ for $i = 1, 2$, where $X_{1,k} = \sqrt{P_1}$ and $X_{2,k} = \sqrt{P_2} \exp(j2\pi \frac{\omega k}{N})$ for $\omega \in \{\tilde{L}_{\max}, \tilde{L}_{\max} + 1, \dots, N - \tilde{L}_{\max}\}$. The equivalent time-domain representation is given by

$$\begin{cases} x_{1,n} = \sqrt{P_1 N} \delta_{n,0}, \\ x_{2,n} = \sqrt{P_2 N} \delta_{\text{mod}(n+\omega, N), 0} \end{cases} \quad (30)$$

for $n = 0, 1, \dots, N - 1$. In [11], based on the Zadoff-Chu sequence [20]

$$z_{n,M} = \begin{cases} \exp\left\{-j\pi \frac{n(n+2q)U}{M}\right\}, & \text{for } M \text{ even} \\ \exp\left\{-j\pi \frac{n(n+1+2q)U}{M}\right\}, & \text{for } M \text{ odd,} \end{cases} \quad (31)$$

where q and U are integers relatively prime to M , the sequences

$$\begin{cases} x_{1,n} = \sqrt{P_1} z_{n, \frac{N}{2}}, \\ x_{2,n} = \sqrt{P_2} z_{n, \frac{N}{2}} \exp(j2\pi \frac{n}{N}) \end{cases} \quad (32)$$

have been proposed: The sequences (32) satisfy the optimal training conditions (24) and (25) if $N > 2\tilde{L}_{\max}$.

Let us now consider the PAPR of a training sequence defined as

$$\text{PAPR} = \frac{\max_n |x_{i,n}|^2}{\frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2}. \quad (33)$$

The training sequences in (30) are simple to generate, but unfortunately, their PAPR is the worst (that is, of value N) since $\max_n |x_{i,n}|^2 = NP_i$ and $\frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2 = P_i$. The high PAPR either increases the design cost or degrades the performance at a similar design cost as delineated in Section 6. In the meantime, the training sequences in (32) have a higher complexity with complex values, but their PAPR is the best (that is, of value 1) since $\max_n |x_{i,n}|^2 = P_i$

and $\frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2 = P_i$.

Taking the PAPR also into consideration, we now propose some sets of optimal training sequences with lower or similar complexity than the conventional training sequences. From the conditions (24) and (25) expressed in the time domain, it is clear that the set of zero-correlation zone (ZCZ) sequences [21,22] with family size 2, sequence length N , and ZCZ \tilde{L}_{\max} can be used as the optimal

training sequences of the LS and LMMSE estimators. Specifically, consider the set

$$\begin{cases} \mathbf{x}_1 = \sqrt{P_1} \begin{bmatrix} -\mathbf{u}^{(n_o)} & \mathbf{v}^{(n_o)} \end{bmatrix}^T, \\ \mathbf{x}_2 = \sqrt{P_2} \begin{bmatrix} -\zeta(\mathbf{u}^{(n_o)}) & -\zeta(\mathbf{v}^{(n_o)}) \end{bmatrix}^T \end{cases} \quad (34)$$

of two binary training sequences with length $N = 2^{n_o+1}$, where $\zeta(\mathbf{a})$ denotes the reverse sequence of \mathbf{a} , and the sequences $\mathbf{u}^{(m)}$ and $\mathbf{v}^{(m)}$, both of length 2^{m+1} , are defined recursively by $\mathbf{u}^{(m)} = [\mathbf{u}^{(m-1)} \ \mathbf{v}^{(m-1)}]$ and $\mathbf{v}^{(m)} = [-\mathbf{u}^{(m-1)} \ \mathbf{v}^{(m-1)}]$ with the initial condition $\mathbf{u}^{(0)} = \mathbf{v}^{(0)} = 1$ [21]. The sequences in (34) satisfy the optimal training condition if $N > 4\tilde{L}_{\max}$.

In addition, we can independently design polyphase [23,24] training sequences satisfying the optimal training condition based on the Chu sequence [25]. Specifically, for a positive integer ν of which the greatest common divisor with N is 1, consider the sequences designed as

$$\begin{cases} x_{1,n} = \sqrt{P_1} \exp(j\pi \frac{\nu}{N} n^2), \\ x_{2,n} = \sqrt{P_2} \exp(j\pi \frac{\nu}{N} (n + \frac{N}{2})^2) \end{cases} \quad (35)$$

for $n = 0, 1, \dots, N - 1$. Since $\sum_{n=0}^{N-1} x_{i,n} x_{i,\text{mod}(n+l, N)}^* = 0$ for $l \neq 0$ and $\sum_{n=0}^{N-1} x_{1,n} x_{2,\text{mod}(n+l, N)}^* = 0$ for $|l| = 0, 1, \dots, \frac{N}{2} - 1$, the sequences in (35) satisfy the conditions (24) and (25) if $N > 2\tilde{L}_{\max}$. Clearly, the PAPR of the proposed training sequences in (34) and (35) is the best (that is, of value 1)

since $\max_n |x_{i,n}|^2 = P_i$ and $\frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2 = P_i$.

In Table 1, we have compared the optimal training sequences discussed so far, where (30) and (32) are the optimal sequences for the LS estimator considered in [10,11], respectively, and (34) and (35) are the optimal sequences proposed in this paper for the LMMSE estimator. For reference, we have also included the raw cubic metric (RCM) of the training sequences. Proposed as an alternative of the PAPR for better estimation on the non-linear effect of a high-power amplifier [26], the RCM of a training sequence $\mathbf{x}_i = [x_{i,0} \ x_{i,1} \ \dots \ x_{i,N-1}]^T$ is computed as

$$\text{RCM} = 20 \log_{10} \sqrt{\frac{\frac{1}{N} \sum_{n=0}^{N-1} (|x_{i,n}|^3)^2}{\left(\frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2\right)^3}} \text{ [dB]}. \quad (36)$$

The RCM is 0 dB when there exists no signal fluctuation and gets larger when the fluctuation in the amplitudes gets more severe. Specifically, the RCMs of the sequences (32), (34), and (35) are all 0 dB since the sequences have

Table 1 Comparison of some optimal training sequences

Sequences	(30) [10]	(32) [11]	(34) [Proposed]	(35) [Proposed]
MSE	Optimal	Optimal	Optimal	Optimal
PAPR	N	1	1	1
RCM	$20 \log_{10} N$	0	0	0
Complexity	Low	High	Low	High

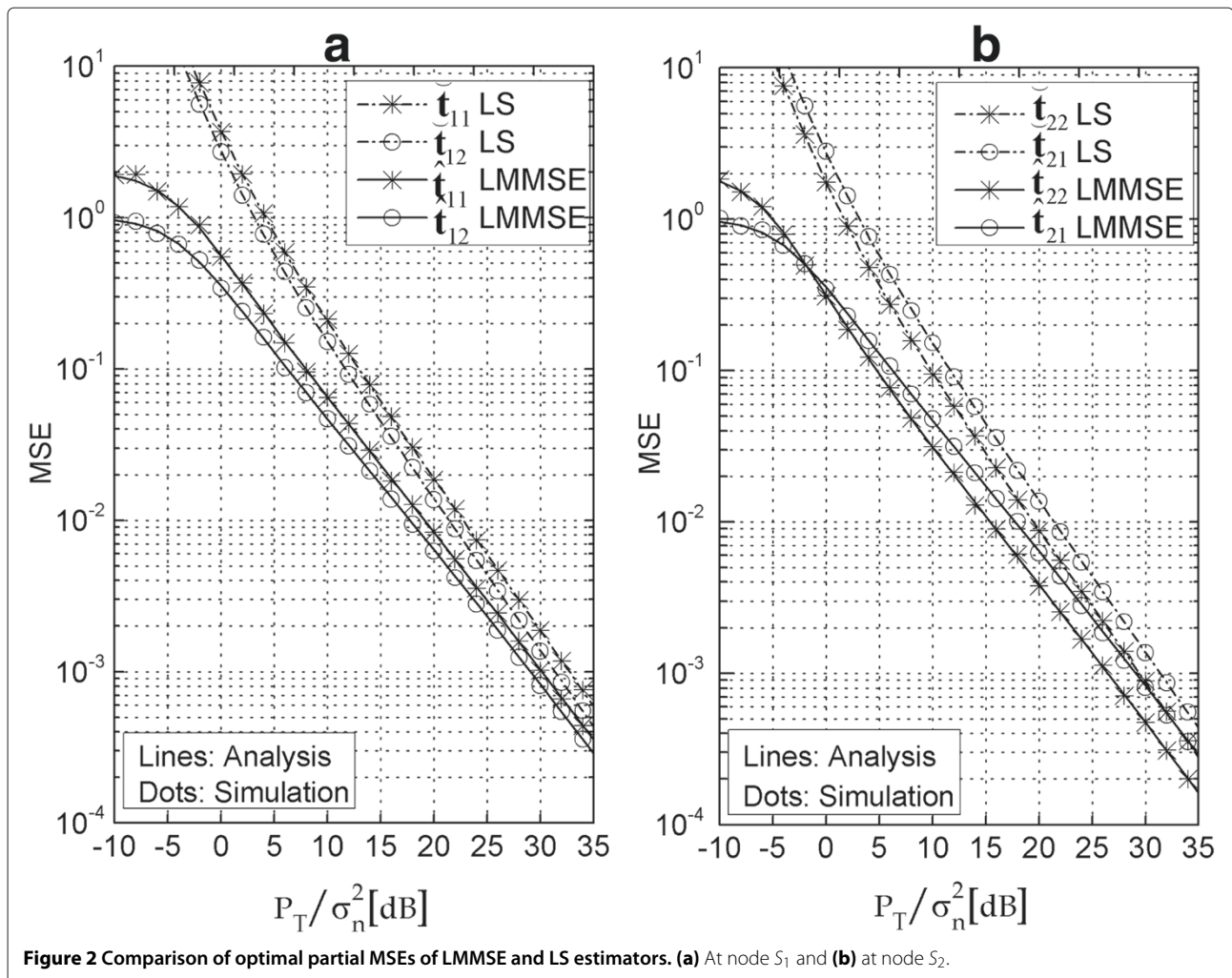
the property of constant modulus $|x_{i,n}| = \sqrt{P_1}$ for $n = 0, 1, \dots, N-1$.

Clearly, the four training sequences all provide the same optimal MSE performance for both the LS and LMMSE estimators. Although the sequences in (30) require the lowest complexity in the memory size and in computation with only a single non-zero value multiplication for N OFDM signal samples, they result in the highest PAPR and RCM, which makes the sequences rather impractical. Among the other three optimal training sequences, all providing the minimum PAPR and RCM, the proposed

sequences in (34), for which a binary value is used for each OFDM sample, have the lowest complexity: A binary value requires only 1-bit storage in the memory, while a complex value in (32) and (35) requires two floating point storage in the memory ($2q$ bits with 2^q quantization levels for each floating point). Therefore, we believe that the proposed sequences in (34) are the most desirable in terms of performance and complexity.

6 Performance evaluation

Let us now evaluate the performance of the LMMSE and LS channel estimators when the total power dissipated by the system is set to P_T and the relay node is located at the center point of the two source nodes. Then, the average power of source nodes is $P_1 = P_2 = \frac{1}{4}P_T$, and that of the relay is $P_r = \frac{1}{2}P_T$. Other system parameters employed in the evaluation of performance comparisons are as follows: Sampling time $T_s = 50$ ns, DFT size $N = 64$, and length of a CP $G = 16$. In the simulation, we have generated 10^5 packets for each source node, where one packet consists



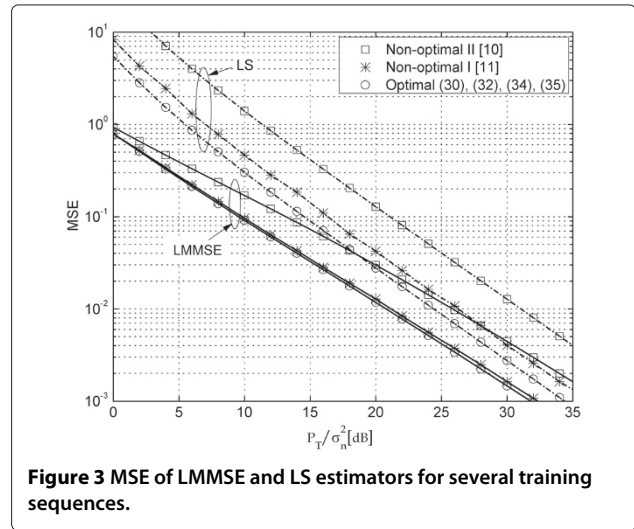
of one OFDM preamble and 49 OFDM data symbols of the same power. The channels remain unchanged over one packet transmission but undergo independent fading from packet to packet. To simulate frequency-selective fading channels, we adopt the exponential MIP

$$\sigma_{h_{i,l}}^2 = \frac{1 - e^{-a_i}}{1 - e^{-a_i L_i}} e^{-a_i l} \quad (37)$$

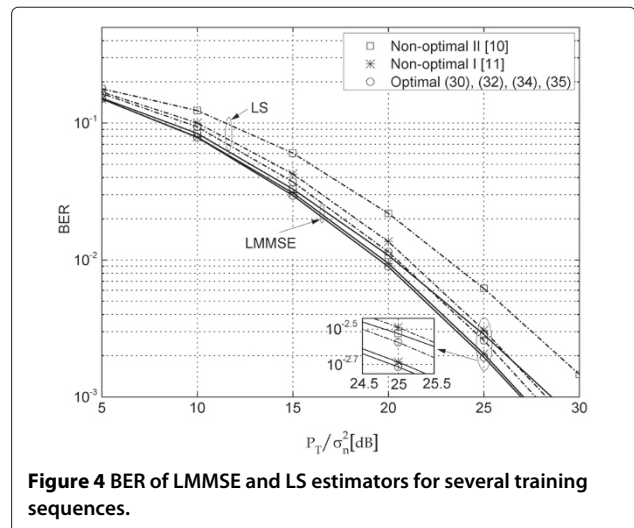
for $l = 0, 1, \dots, L_i - 1$ with $a_i = \frac{10}{L_i - 1}$. This MIP results in $\beta_1 = \beta_2 = 1$. In this setup, the MSE is computed with 10^5 OFDM preambles, while the bit error rate (BER) is calculated with 49×10^5 OFDM data symbols using BPSK modulation.

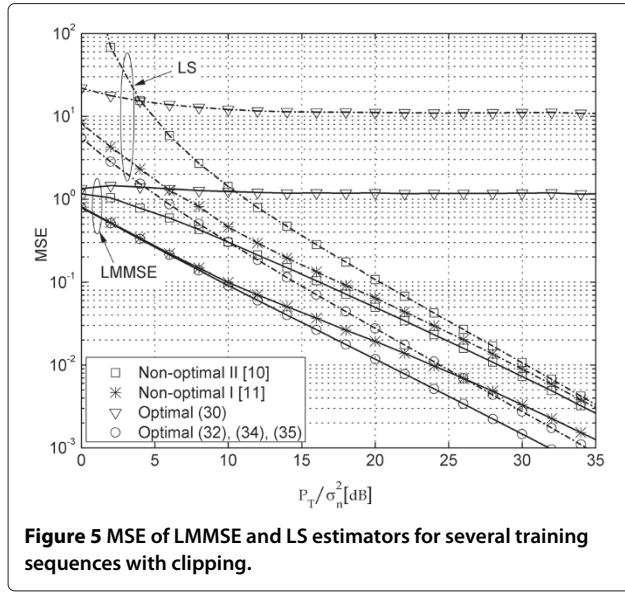
Figure 2 verifies our analysis by comparing the optimal partial MSEs (27) and (46) of the LMMSE and LS estimators, respectively, with simulation results when the optimal training sequences in (35) are employed and the channels have the exponential MIP (37) with $L_1 = 8$ and $L_2 = 4$. Figure 2a,b compares the results at S_1 and S_2 , respectively. It is clearly observed that the results from the analysis agree well with those from the simulation. In addition, by mitigating the noise enhancement incurred in the LS estimator, the LMMSE estimator performs better than the LS estimator with the gain higher in the lower SNR region. The gain diminishes as the SNR increases since the effect of the noise components is reduced. In the high SNR region, the partial MSE is the largest and smallest in estimating t_{11} and t_{22} , respectively, for both the LMMSE and LS estimators since the partial MSE in estimating t_{im} is proportional to $\tilde{L}_{im} = L_i + L_m - 1$ in the high SNR region. Specifically, with $\tilde{L}_{11} = 2 \cdot 8 - 1 = 15$, $\tilde{L}_{12} = \tilde{L}_{21} = 8 + 4 - 1 = 11$, and $\tilde{L}_{22} = 2 \cdot 4 - 1 = 7$ in (28) and (46), we have $\frac{\text{MSE}(\hat{t}_{21})}{\text{MSE}(\hat{t}_{22})} \approx \frac{\text{MSE}(\check{t}_{21})}{\text{MSE}(\check{t}_{22})} = \frac{11}{7}$ and $\frac{\text{MSE}(\hat{t}_{11})}{\text{MSE}(\hat{t}_{22})} \approx \frac{\text{MSE}(\check{t}_{11})}{\text{MSE}(\check{t}_{22})} = \frac{15}{7}$. For the LMMSE estimator, the partial MSE of \hat{t}_{ii} is observed to be slightly larger than the partial MSE of \hat{t}_{ij} for $(i, j) \in \{(1, 2), (2, 1)\}$ when the SNR is very low due to the factor $1 + \delta_{i,i} = 2$ (on the RHS of (22)) in the trace of $\mathbf{R}_{t_{ii}}$, which has eventually led to the factor $1 + \delta_{i,i} = 2$ multiplied to $\beta_i \beta_m$ when $m = i$ (on the RHS of (29)). Note that t_{ii} , a convolution of the identical CIR \mathbf{h}_i and \mathbf{h}_i , exhibits higher correlation among the elements than t_{ij} , a convolution of two independent CIRs \mathbf{h}_i and \mathbf{h}_j .

Figures 3 and 4 compare the MSEs and BERs, respectively, of the LMMSE and LS estimators when optimal and non-optimal training sequences are used, where we have assumed the symmetric channels with $L_1 = L_2 = 6$ for the exponential MIP, leading to an identical MSE (BER) performance at S_1 and S_2 . The MSE in Figure 3 represents $\text{MSE}(\hat{t}_1)$ and $\text{MSE}(\check{t}_1)$ of the LMMSE and LS estimators, respectively, and the BER in Figure 4 is evaluated at S_1 . In these figures, ‘optimal (30), (32), (34), (35)’ denotes the results with the optimal training sequences in (30),



(32), (34), and (35)^b, ‘Non-optimal I’ denotes the results with the training sequences constructed with randomly generated BPSK symbols $\{X_{i,k}\}$ used as a benchmark in [11], and ‘Non-optimal II’ denotes the results with the training sequence $x_{i,n} = \sqrt{0.1}x_{i,n}^*$ for $0 \leq n \leq 31$ and $x_{i,n} = \sqrt{1.9}x_{i,n}^*$ for $32 \leq n \leq 63$ used as a benchmark in [10]. It is clearly confirmed in these figures that all the optimal training sequences (conventional or proposed) provide the best performance for both the LMMSE and LS estimators. It is also observed that non-optimal sequences could cause significant degradation in the performance of channel estimation. Although the sequences ‘Non-optimal I’ provide a similar performance with the optimal training sequence, the former exhibits a higher PAPR than the latter. Interestingly, it is observed that the performance degradation incurred by using non-optimal

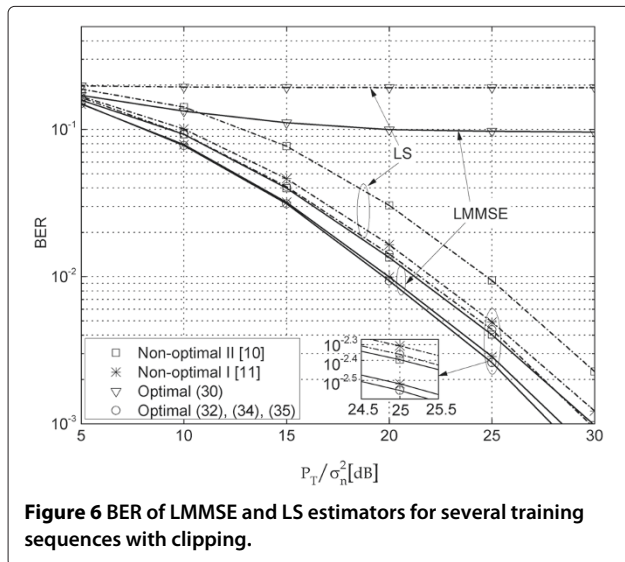




training sequences is smaller in the LMMSE estimation than in the LS estimation.

To incorporate a partial effect of the PAPR of transmitted signals, the MSEs and BERs of Figures 3 and 4 are now re-evaluated in Figures 5 and 6 when there exists a baseband non-linearity at the source nodes. We model the baseband non-linearity by a hard limiter as in [27], where the transmitted signals $\{x_{i,n}\}$ are clipped when $|x_{i,n}| \geq A$.

The clipping ratio defined by $\frac{A}{\sigma_{x_i}}$ with $\sigma_{x_i}^2 = \frac{1}{N} \sum_{n=0}^{N-1} |x_{i,n}|^2$ is assumed to be 6 dB. It is clearly observed that the optimal training sequences in (32), (34), and (35) are not influenced by the clipping, while the conventional optimal training sequences in (30), due to its high PAPR, are not



appropriate in practical systems. In addition, the proposed sequences in (34) and (35) can provide the same performance at a lower or similar complexity when compared with the conventional optimal sequences in (32).

7 Conclusion

In this paper, we have addressed the LMMSE estimation of channels for OFDM-based TWR systems employing ANC. We have derived the conditions for optimal training sequences minimizing the MSE of the LMMSE estimator, which is found out to be equivalent to the optimal training conditions of the LS estimator. In addition, the MSE behavior of the LMMSE estimator with optimal training sequences is analyzed to be generally dependent on the MIP of the channel. Results from analysis and simulation show that the LMMSE estimator outperforms the LS estimator at most practical values of SNR.

Taking the MSE and PAPR into account simultaneously, we have suggested new designs of optimal training sequences exhibiting lower or similar complexity when compared with the conventional optimal training sequences. Simulation results also show that the proposed training sequences would perform adequately even when signal clipping occurs during signal processing.

Endnotes

^aFor two vectors $\mathbf{a} = [a_0 \ a_1 \ \dots \ a_{N-1}]^T$ and $\mathbf{b} = [b_0 \ b_1 \ \dots \ b_{N-1}]^T$, the l th element of $\mathbf{c} = \mathbf{a} \otimes \mathbf{b}$ is given by $c_l = \sum_{n=0}^{N-1} a_{n-l} b_{\text{mod}(n,N)} = \sum_{n=0}^{N-1} a_n b_{\text{mod}(n+l,N)}$, where $\text{mod}(x, y)$ denotes the remainder after dividing x by y .

^bThe MSEs of the optimal training sequences are all the same, and thus, we have shown only the MSE of (34) to make the graphs less crowded.

Appendix 1

Proof of Theorem 1

Let $\mathbf{d} = [d_0 \ d_1 \ \dots \ d_{\tilde{L}_i-1}]$ and $\boldsymbol{\lambda} = [\lambda_0 \ \lambda_1 \ \dots \ \lambda_{\tilde{L}_i-1}]$ denote the vectors of diagonal elements and eigenvalues, respectively, of $\boldsymbol{\Upsilon}_i = \mathbf{R}_{t_i}^{-1} + \theta_i \boldsymbol{\Psi}_i^H \boldsymbol{\Psi}_i$. Since the matrix $\boldsymbol{\Upsilon}_i$ is symmetric, the vector \mathbf{d} is majorized by the vector $\boldsymbol{\lambda}$ [28], that is, we have

$$\sum_{k=0}^K d_{[k]} \leq \sum_{k=0}^K \lambda_{[k]} \quad (38)$$

if $K = 0, 1, \dots, \tilde{L}_i - 2$, and

$$\sum_{k=0}^K d_{[k]} = \sum_{k=0}^K \lambda_{[k]} \quad (39)$$

if $K = \tilde{L}_i - 1$, where $d_{[k]}$ and $\lambda_{[k]}$ denote the k th largest elements in $\{d_l\}_{l=0}^{\tilde{L}_i-1}$ and $\{\lambda_l\}_{l=0}^{\tilde{L}_i-1}$, respectively. In addition,

since $\text{tr} \{ \Upsilon_i^{-1} \} = \sum_{k=0}^{\tilde{L}_i-1} \frac{1}{\lambda_k}$ is Shur-convex [29], we have

$$\text{tr} \{ \Upsilon_i^{-1} \} \geq \sum_{k=0}^{\tilde{L}_i-1} \frac{1}{d_k}. \quad (40)$$

Hence, noting that \mathbf{R}_{t_i} is diagonal, the MSE (15) is minimized when $\Upsilon_i = \mathbf{R}_{t_i}^{-1} + \theta_i \Psi_i^H \Psi_i$ is a diagonal matrix, or equivalently, when $\Psi_i^H \Psi_i$ is a diagonal matrix. Thus, since the diagonal elements of $\Psi_{ii}^H \Psi_{ii}$ and $\Psi_{ij}^H \Psi_{ij}$ are identically $\|\mathbf{x}_i\|^2$ and $\|\mathbf{x}_j\|^2$, respectively, and we have $\|\mathbf{x}_i\|^2 = NP_i$ for $i = 1$ and 2 , it is required that

$$\Psi_i^H \Psi_i = \begin{bmatrix} \Psi_{ii}^H \Psi_{ii} & \Psi_{ij}^H \Psi_{ij} \\ \Psi_{ij}^H \Psi_{ii} & \Psi_{ij}^H \Psi_{ij} \end{bmatrix} = \begin{bmatrix} NP_i \mathbf{I}_{\tilde{L}_{ii}} & \mathbf{0}_{\tilde{L}_{ii} \times \tilde{L}_{ij}} \\ \mathbf{0}_{\tilde{L}_{ij} \times \tilde{L}_{ii}} & NP_j \mathbf{I}_{\tilde{L}_{ij}} \end{bmatrix} \quad (41)$$

for $(i, j) \in \{(1, 2), (2, 1)\}$ to minimize the MSE at S_i . The condition (41) is equivalent to (24) and (25).

Appendix 2

The LS estimator and its optimal training condition

With the received signal (8), the LS estimator $\check{\mathbf{t}}_i = \arg \min_{\mathbf{a}} \|\mathbf{y}_i - \alpha \Psi_i \mathbf{a}\|^2$ of \mathbf{t}_i is obtained by setting the first derivative

$$\frac{\partial}{\partial \mathbf{a}} \|\mathbf{y}_i - \alpha \Psi_i \mathbf{a}\|^2 = 2\alpha \Psi_i^H (\mathbf{y}_i - \alpha \Psi_i \mathbf{a}) \quad (42)$$

to zero, as

$$\check{\mathbf{t}}_i = \frac{1}{\alpha} (\Psi_i^H \Psi_i)^{-1} \Psi_i^H \mathbf{y}_i \quad (43)$$

since the second derivative is positive semi-definite, $\frac{\partial}{\partial \mathbf{a}} \left(\frac{\partial}{\partial \mathbf{a}} \|\mathbf{y}_i - \alpha \Psi_i \mathbf{a}\|^2 \right)^H = 2\Psi_i^H \Psi_i \geq 0$. The MSE of the LS estimator is then given by

$$\begin{aligned} \text{MSE}(\check{\mathbf{t}}_i) &= \text{tr} \left\{ \frac{\sigma_n^2}{\alpha^2} (\Psi_i^H \Psi_i)^{-1} \Psi_i^H (\alpha^2 \beta_1 + 1) \Psi_i (\Psi_i^H \Psi_i)^{-1} \right\} \\ &= \sigma_n^2 \left(\beta_1 + \frac{\beta_1 P_1 + \beta_2 P_2 + \sigma_n^2}{P_r} \right) \text{tr} \left\{ (\Psi_i^H \Psi_i)^{-1} \right\}, \end{aligned} \quad (44)$$

which is minimized when $\Psi_i^H \Psi_i$ is diagonal from the same reason as that applied when proving Theorem 1 in Appendix 1. Under the power constraint $\|\mathbf{x}_i\|^2 = NP_i$ for $i = 1, 2$, the optimal training conditions of the LS estimator thus becomes the same as those of the LMMSE estimator given in (41), and the resultant minimum MSE of the LS estimator is given by

$$\text{MSE}_o(\check{\mathbf{t}}_i) = \text{MSE}_o(\check{\mathbf{t}}_{i1}) + \text{MSE}_o(\check{\mathbf{t}}_{i2}), \quad (45)$$

where

$$\text{MSE}_o(\check{\mathbf{t}}_{im}) = \sigma_n^2 \left(\beta_1 + \frac{\beta_1 P_1 + \beta_2 P_2 + \sigma_n^2}{P_r} \right) \frac{\tilde{L}_{im}}{NP_m} \quad (46)$$

is the partial MSE for $m = 1, 2$. We would again like to mention that the optimal MSE (45) of the LS estimator derived herein is identical to that derived in [10,11] and becomes the high SNR approximation (28) of the optimal MSE of the LMMSE estimator.

In [10,11], the optimal training condition (41) is expressed with the frequency-domain representation $\mathbf{X}_i = \mathbf{W} \mathbf{x}_i = [X_{i,0} \ X_{i,1} \ \cdots \ X_{i,N-1}]$ of the training sequences as

$$\mathbf{W}_{\tilde{L}_{im}}^H \text{diag}(\mathbf{X}_i^H \odot \mathbf{X}_i) \mathbf{W}_{\tilde{L}_{im}} = P_i \mathbf{I}_{\tilde{L}_{im}}, \quad (47)$$

$$\mathbf{W}_{\tilde{L}_{im}}^H \text{diag}(\mathbf{X}_i^H \odot \mathbf{X}_j) \mathbf{W}_{\tilde{L}_{mj}} = \mathbf{0}_{\tilde{L}_{im} \times \tilde{L}_{mj}} \quad (48)$$

for $m = i, j$ and $(i, j) = \{(1, 2), (2, 1)\}$, where \odot is the element-wise multiplication. Let us now show that the conditions (47) and (48) are equivalent to those given as (24) and (25). First, (47) can be rewritten as

$$\frac{1}{N} \sum_{k=0}^{N-1} X_{i,k}^* X_{i,k} \exp \left\{ -j2\pi \frac{k(l_2 - l_1)}{N} \right\} = \begin{cases} P_i, & \text{for } l_1 = l_2, \\ 0, & \text{for } l_1 \neq l_2, \end{cases} \quad (49)$$

for $i = 1, 2$, where $l_1, l_2 = 0, 1, \dots, \max(\tilde{L}_{i1}, \tilde{L}_{i2}) - 1$. With $\tilde{L}_i = \max(\tilde{L}_{i1}, \tilde{L}_{i2})$, we can rewrite (49) as

$$\sum_{k=0}^{N-1} |X_{i,k}|^2 \exp \left(-j2\pi \frac{kl}{N} \right) = \begin{cases} NP_i, & \text{for } l = 0, \\ 0, & \text{for } |l| = 1, 2, \dots, \tilde{L}_i - 1 \end{cases} \quad (50)$$

for $i = 1, 2$. Similarly, (48) can be rewritten as

$$\sum_{k=0}^{N-1} X_{1,k} X_{2,k}^* \exp \left\{ -j2\pi \frac{k(l_1 - l_2)}{N} \right\} = 0 \quad (51)$$

for $l_1, l_2 = 0, 1, \dots, \max(\tilde{L}_1, \tilde{L}_2) - 1$, which is equivalent to

$$\sum_{k=0}^{N-1} X_{1,k}^* X_{2,k} \exp \left(-j2\pi \frac{kl}{N} \right) = 0 \quad (52)$$

for $|l| = 0, 1, \dots, \tilde{L}_{\max} - 1$. Since

$$\sum_{k=0}^{N-1} X_{i,k}^* X_{m,k} \exp \left(-j2\pi \frac{kl}{N} \right) = \sum_{n=0}^{N-1} x_{i,n}^* x_{m,\text{mod}(n+l,N)} \quad (53)$$

for $m = i, j$ from the (generalized) Parseval's theorem [30], it is straightforward to see that (47) and (48) represented in the frequency-domain are equivalent to (24) and (25), respectively, represented in the time domain.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by the National Research Foundation of Korea, with funding from the Ministry of Education, Science, and Technology, under grants 2012-0001867, 2012R1A1A2040091, and 2012-0005622.

Author details

¹Department of Electronics and Radio Engineering, Kyung Hee University, Yongin Gyeonggi-do 446-701, Korea. ²Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea.

Received: 29 October 2012 Accepted: 15 May 2013

Published: 28 May 2013

References

1. JN Laneman, GW Wornell, Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inform. Theory* **49**(10), 2415–2425 (2003)
2. JN Laneman, DNC Tse, GW Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inform. Theory* **50**(12), 3062–3080 (2004)
3. S Berger, M Kuhn, A Wittneben, Recent advances in amplify-and-forward two-hop relaying. *IEEE Comm. Mag.* **47**(7), 50–56 (2009)
4. P Popovski, H Yomo, Wireless network coding by amplify-and-forward for bi-directional traffic flows. *IEEE Comm. Lett.* **11**(1), 16–18 (2007)
5. B Rankov, A Wittneben, Spectral efficient protocols for half-duplex fading relay channels. *IEEE J. Select. Areas Comm.* **25**(2), 379–389 (2007)
6. C Ho, R Zhang, YC Liang, in *Proc. IEEE Int. Conf. Comm. (ICC)*. Two-way relaying over OFDM: optimized tone permutation and power allocation (Beijing, China, May 2008)
7. Z Li, X Xia, B Li, Achieving full diversity and fast ML decoding via simple analog network coding for asynchronous two-way relay networks. *IEEE Trans. Comm.* **57**(12), 3672–3671 (2009)
8. HN Vu, HY Kong, Joint subcarrier matching and power allocation in OFDM two-way relay systems. *J. Comm. Net.* **14**(3), 257–266 (2012)
9. F Gao, R Zhang, YC Liang, Optimal channel estimation and training design for two-way relay networks. *IEEE Trans. Comm.* **57**(10), 3024–3033 (2009)
10. F Gao, R Zhang, YC Liang, Channel estimation for OFDM modulated two-way relay networks. *IEEE Trans. Signal Process.* **57**(11), 4443–4455 (2009)
11. W Yang, Y Cai, J Hu, W Yang, Channel estimation for two-way relay OFDM networks. *EURASIP J. Wireless Comm. Network.* **2010**(186182), 1–6 (2010)
12. T Pham, YC Liang, A Nallanathan, H Garg, Optimal training sequences for channel estimation in bi-directional relay networks with multiple antennas. *IEEE Trans. Comm.* **58**(2), 474–479 (2010)
13. G Wang, F Gao, Y Wu, C Tellambura, Joint CFO and channel estimation for OFDM-based two-way relay networks. *IEEE Trans. Wireless Comm.* **10**(2), 456–465 (2011)
14. T Jiang, Y Wu, An overview: peak-to-average power ratio reduction techniques for OFDM signals. *IEEE Trans. Broadcast.* **54**(2), 257–268 (2008)
15. H Gacanin, T Sjödin, F Adachi, On channel estimation for analog network coding in a frequency-selective fading channel. *EURASIP J. Wireless Comm. Network.* **2011**: 980430, 1–12 (2011)
16. I Prodan, T Obara, F Adachi, H Gacanin, Performance of pilot-assisted channel estimation without feedback for broadband ANC systems using OFDM access. *EURASIP J. Wireless Comm. Netw.* **2012**: 315, 1–21 (2012)
17. YG Li, JH Winters, NR Sollenberger, Simplified channel estimation for OFDM systems with multiple transmit antennas. *IEEE Trans. Wireless Comm.* **1**(1), 67–75 (1999)
18. KJ Lee, IK Lee, in *Proc. IEEE Veh. Technol. Conf. (VTC-Spring)*. Achievable rate regions for two-way MIMO AF multiple-relay channels (Budapest, Hungary, May 2011)
19. LL Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Addison-Wesley, Reading, 1991)
20. BM Popović, Generalized chirp-like polyphase sequences with optimum correlation properties. *IEEE Trans. Inform. Theory*, **38**(4), 1406–1409 (1992)
21. PZ Fan, N Suehiro, N Kuroyanagi, XM Deng, Class of binary sequences with zero correlation zone. *Electron. Lett.* **35**(10), 777–779 (1999)
22. H Torri, M Nakamura, N Suehiro, A new class of zero-correlation zone sequences. *IEEE Trans. Inform. Theory*, **50**(3), 559–565 (2004)
23. SI Park, SR Park, I Song, N Suehiro, Multiple-access interference reduction for QS-CDMA systems with a novel class of polyphase sequences. *IEEE Trans. Inform. Theory*, **46**(4), 1448–1458 (2000)
24. SR Park, I Song, S Yoon, J Lee, A new polyphase sequence with perfect even and good odd cross correlation functions for DS/CDMA systems. *IEEE Trans. Vehic. Techn.* **51**(5), 855–866 (2002)
25. DC Chu, Polyphase codes with good periodic correlation properties. *IEEE Trans. Inform. Theory*, **18**(4), 531–532 (1972)
26. Motorola, Cubic metric in 3GPP LTE. The 3rd Generation Partnership Project (3GPP) TSG-RAN WG1 Meeting, Tdoc# R1-060023, Helsinki, Finland, Jan. 2006
27. L Wang, C Tellambura, A simplified clipping and filtering technique for PAR reduction in OFDM systems. *IEEE Sig. Process. Lett.* **12**(5), 453–456 (2005)
28. AW Marshall, I Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic, New York, 1979)
29. JH Kotecha, AM Sayeed, Transmit signal design for optimal estimation of correlated MIMO channels. *IEEE Trans. Signal Process.* **52**(2), 546–557 (2004)
30. AV Oppenheim, RW Schaffer, *Discrete-Time Signal Processing*, 2nd Ed. (Prentice Hall, Upper Saddle River, 1999)

doi:10.1186/1687-1499-2013-140

Cite this article as: Tran et al.: Channel estimation and optimal training with the LMMSE criterion for OFDM-based two-way relay networks. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:140.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com