Two-Way Discriminatory Channel Estimation for Non-Reciprocal Wireless MIMO Channels

Chao-Wei Huang[⋆], Tsung-Hui Chang[⋆], Xiangyun Zhou[†] and Y.-W. Peter Hong[⋆]

[⋆]Institute of Commun. Eng. & Department of Elect. Eng., National Tsing Hua University, Hsinchu, Taiwan

[†]Research School of Engineering, The Australian National University, Canberra, Australia

Emails: cwhuang@erdos.ee.nthu.edu.tw, tsunghui.chang@ieee.org, xiangyun.zhou@anu.edu.au, ywhong@ee.nthu.edu.tw

***Abstract*—The idea of discriminatory channel estimation (DCE) was recently proposed to differentiate the performance at a legitimate receiver (LR) and an unauthorized receiver (UR) by carefully designing the training phase of transmission. In particular, an artificial noise (AN) signal is inserted into the training signal to degrade the channel estimation performance at UR. However, without careful design, the inserted AN may potentially leak into LR's channel causing degradation in LR's performance. In order to minimize such an undesirable AN leakage, the transmitter must acquire sufficient knowledge of the downlink channel to LR for careful signalling, which is particularly difficult to achieve in non-reciprocal channel environments. In this work, we design a DCE scheme with multiple training stages to achieve this task. The key feature of the proposed scheme is to utilize a two-way training scheme that allows both the transmitter and LR to collaboratively send different training signals to achieve DCE. Simulation results demonstrate that our proposed scheme is able to achieve accurate channel estimation and data detection performances at LR, while severely degrading UR's performances.**

I. Introduction

Due to the broadcast nature of the wireless medium, eavesdropping or overhearing by unauthorized receivers has become much easier compared to conventional wireline systems and, thus, maintaining secrecy in wireless transmissions has become an increasingly important and challenging problem. In the past, cryptographic encryption in the application layer has been widely used in solving these issues. On the other hand, the idea of physical layer security, which provides an alternative or complement to achieve secrecy through coding and signal processing in the physical layer, has been extensively studied recently.

Most work on the physical layer security focuses on the data transmission phase and assumes perfect channel training. However, channel estimation is never perfect in reality and data detection performance considerably counts on the accuracy of the channel estimate for coherent detection. This fact suggests that secrecy can be achieved by discriminating channel estimation performances between the legitimate receiver (LR) and the unauthorized receiver (UR), which is referred to as discriminatory channel estimation (DCE). The idea of DCE was first addressed in [1] where the artificial noise (AN) [2] is inserted on top of the training signal to degrade the UR's channel estimation performance. However, the AN must be placed in the null space of transmitter-to-LR channel to minimize the interference on the LR. This requires the transmitter having the knowledge of the LR's channel which is achieved through feedback from the LR. In the original DCE scheme, only the transmitter is allowed to transmit the training signal. Moreover, the training signal sending from the transmitter benefits both the LR and UR. To discriminate the channel estimation performance between the LR and UR, multiple stages of feedback and retraining are needed which results in significant training overhead.

In this paper, we propose an efficient DCE scheme using the two-way training methodology which allows both the transmitter and receiver to send training signals. The idea of two-way training schemes were originally studied for conventional point-to-point links (without secrecy considerations) in [3]–[5] to obtain the CSI at both the receiver and the transmitter without the use of feedback. We adopt the concept of two-way training into the DCE design. Our previous work in [6] presented a two-way DCE scheme for the case of reciprocal channels. Due to the channel reciprocity, a single reverse training stage would be sufficient for the transmitter to estimate the LR's channel. In this work, we focus on a more general and complex scenario of non-reciprocal channels, which is common, e.g., in frequency-division multiplexing (FDD) systems. In this case, an additional training stage, called round-trip training, is needed where the transmitter first broadcasts a randomly generated signal, which is then echoed back from the LR to the transmitter. The echoed signal contains information of both the transmitter-to-LR (downlink) and LR-to-transmitter (uplink) channels. The uplink channel can be estimated by the transmitter using the reverse training. Using the received signals from both the round-trip training and the reverse training, the transmitter is able to estimate the desired downlink channel. Compared to the multi-stage feedback-and-retraining DCE scheme in [1], the newly proposed two-way training scheme drastically decreases the overall training overhead and design complexity. Our simulation result shows that the proposed DCE design can efficiently discriminate the performance of channel estimation and the data detection between the LR and UR.

Notations: Boldface is used for matrices $\mathbf{X}$ (upper case) and vectors $\mathbf{x}$ (lower case). $\mathbf{X}^T$, $\mathbf{X}^*$ and $\mathbf{X}^H$ denote the transpose,

This work was supported by in part by the National Science Council, Taiwan, under grant NSC 100-2628-E-007-025-MY3 and by the Australian Research Council's Discovery Projects funding scheme (project no. DP110102548).

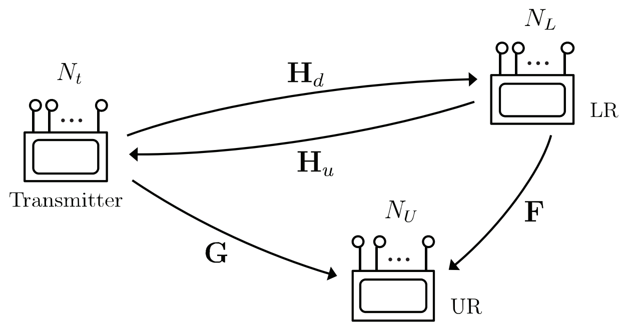

Fig. 1: A wireless MIMO system consisting of a transmitter, a legitimate receiver (LR) and an unauthorized receiver (UR).

the complex conjugate and the Hermitian of the matrix $\mathbf{X}$, respectively. Let $\mathbf{0}_{M\times N}$ be the M by N zero matrix and $\mathbf{I}_M$ be the M by M identity matrix. Tr$(\cdot)$ denotes the trace of a square matrix and vec$(\cdot)$ is the operator which stacks the columns of a matrix into a vector. The symbol $\otimes$ denotes the Kronecker matrix product, and E$\{\cdot\}$ denotes the expectation operator.

II. System Model

Consider a FDD wireless MIMO system which consists of a transmitter, a LR and an UR, as shown in Fig. 1. We assume that the transmitter, the LR and the UR are respectively outfitted with N_t, N_L and N_U antennas, with $N_t > N_L$. The channels between any two of them are quasi-static within one transmission block which has a training phase and a data transmission phase. The wireless channels are non-reciprocal, that is, the downlink channel (from the transmitter to the LR) is in general different from the uplink channel (from the LR to the transmitter). We denote the downlink channel as $\mathbf{H}_d \in \mathbb{C}^{N_t\times N_L}$ and the uplink channel as $\mathbf{H}_u \in \mathbb{C}^{N_L\times N_t}$ to represent their heterology. Each element of $\mathbf{H}_d$ and $\mathbf{H}_u$ are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variances respectively being $\sigma^2_{H_d}$ and $\sigma^2_{H_u}$. On the other hand, let the channel from the transmitter to the UR be denoted by $\mathbf{G} \in \mathbb{C}^{N_t\times N_U}$ of which each entry is also *i.i.d.* Gaussian random variables with zero mean and variance σ^2_G.

The UR is assumed to passively intercept the message sent by the transmitter. From the legitimate user's viewpoint, the message sent by the transmitter should only be accurately obtained by the LR but not the UR. To reduce the possibility of interception by the UR, we propose a two-way training design for DCE which allows the LR to achieve a good channel estimation performance while confining the UR's channel estimation performance to a poor level. In non-reciprocal channel environments, the two-way DCE scheme consists of three training stages which are a round-trip training, a reverse training and a forward training. The first two stages facilitate the transmitter to obtain knowledge of the downlink channel so that in the following third stage, *i.e.,* the forward training, this channel information can be used in designing the forward training signal for discriminating the channel estimation performance between the LR and UR. The proposed DCE scheme is described in detail in the next section.

III. Proposed Two-Way DCE Design

A. Two-Way Training Strategy

The proposed two-way DCE is achieved by the following three-step training process.

Round-trip Training: First, the transmitter broadcasts a secret signal randomly generated by itself, which is unknown to both the LR and UR. Upon receiving this signal, the LR feeds it back to the transmitter using a simple amplify-and-forward protocol. The go-and-back process makes this stage be referred to as the round-trip training. The received signal of the transmitter, which is called the echoed signal contains information combining the downlink and uplink channels. If the knowledge of uplink channel is available at the transmitter, we have the ability to extract the information about the downlink channel.

Specifically, the secret signal sent by the transmitter is given by

$$\mathbf{X}_{t0} = \sqrt{\frac{\mathcal{P}_0\tau_0}{N_t}}\mathbf{C}_{t0} \tag{1}$$

where $\mathbf{C}_{t0} \in \mathbb{C}^{\tau_0\times N_t}$ is a random matrix satisfying Tr$(\mathbf{C}^H_{t0}\mathbf{C}_{t0}) = N_t$, and $\mathcal{P}_0$ and τ_0 are the transmission power and training length, respectively. The received signal at the LR is given by

$$\mathbf{Y}_{L0} = \mathbf{X}_{t0}\mathbf{H}_d + \mathbf{W}_0 \tag{2}$$

where $\mathbf{W}_0$ is the additive white Gaussian noise (AWGN) matrix at the LR with each element being zero mean and variance σ^2_w. Then, LR amplifies and forwards its received signal back to the transmitter. The echoed signal received at the transmitter side is given by

$$\mathbf{Y}_{t1} = \alpha(\mathbf{X}_{t0}\mathbf{H}_d + \mathbf{W}_0)\mathbf{H}_u + \widetilde{\mathbf{W}}_1 \tag{3}$$

where $\widetilde{\mathbf{W}}_1$ is the AWGN matrix with each entry being zero mean and variance $\sigma^2_{\tilde{w}}$ and the amplifying gain at the LR is given by

$$\alpha = \sqrt{\frac{\mathcal{P}_1\tau_0}{\mathcal{P}_0\tau_0 N_L\sigma^2_{H_d} + \tau_0 N_L\sigma^2_w}} \tag{4}$$

where $\mathcal{P}_1$ is the transmission power of the LR.

Reverse Training: The LR then transmits a training signal to enable the uplink channel estimation at the transmitter. Specifically, the reverse training is given by

$$\mathbf{X}_{L2} = \sqrt{\frac{\mathcal{P}_2\tau_2}{N_L}}\mathbf{C}_{L2} \tag{5}$$

where $\mathbf{C}_{L2} \in \mathbb{C}^{\tau_2\times N_L}$ is the pilot matrix satisfying $\mathbf{C}^H_{L2}\mathbf{C}_{L2} = \mathbf{I}_{N_L}$. Note that this orthogonal design of pilot matrix was shown to be optimal in [7]. Also, $\mathcal{P}_2$ and τ_2 are the transmission power and training length of the LR, respectively. Then, the received signal at the transmitter is given by

$$\mathbf{Y}_{t2} = \mathbf{X}_{L2}\mathbf{H}_u + \widetilde{\mathbf{W}}_2 \tag{6}$$

$$\mathrm{NMSE}_L \approx \left(\frac{1}{\sigma^2_{H_d}} + \frac{\mathcal{P}_3\tau_3}{N_t} \frac{1}{(N_t - N_L)\sigma_a^2 \left(\sigma^2_{H_d} - \sigma^2_{H_d} \frac{\sigma^2_{H_d}\mathcal{P}_0\tau_0}{\sigma^2_{H_d}\mathcal{P}_0\tau_0 + N_t\sigma_w^2} \frac{N_t\sigma^2}{\beta + N_t\sigma^2} \right) + \sigma_w^2} \right)^{-1} \tag{14}$$

where

$$\sigma^2 = \frac{\sigma^4_{H_u}\mathcal{P}_2\tau_2}{\sigma^2_{H_u}\mathcal{P}_2\tau_2 + N_L\sigma^2_{\tilde{w}}} \tag{15}$$

and

$$\beta = N_L \left(\frac{1}{\sigma^2_{H_u}} + \frac{\mathcal{P}_2\tau_2}{N_L\sigma^2_{\tilde{w}}} \right)^{-1} + \frac{\sigma^2_{\tilde{w}}}{\alpha^2\sigma^2_{H_d}\sigma^2_w} \left(\frac{1}{\sigma^2_{H_d}} + \frac{\mathcal{P}_0\tau_0}{N_t\sigma^2_w} \right)^{-1}. \tag{16}$$

where $\widetilde{\mathbf{W}}_2$ is the AWGN matrix with noise power $\sigma^2_{\tilde{w}}$. As the uplink channel estimate is available at the transmitter, with the help of the echoed signal, the transmitter is able to obtain the downlink channel estimate, which is denoted by $\widehat{\mathbf{H}}_{d,t}$. The detailed derivation of downlink channel estimate is shown in Appendix. Note that the knowledge of the reverse training pilot matrix is useless to the UR, since it only helps the UR to estimate the LR-UR channel which is irrelevant to data transmission.

Forward Training: In the forward training stage, the transmitter adds the AN into the training signal to disrupt the UR's channel estimation performance. Meanwhile, having the downlink channel estimate at hand, the AN is placed in the left null space of $\widehat{\mathbf{H}}_{d,t}$ to minimize its interference to the LR. Specifically, with the assumption that $N_t > N_L$, the forward training signal is given by

$$\mathbf{X}_{t3} = \sqrt{\frac{\mathcal{P}_3\tau_3}{N_t}}\mathbf{C}_{t3} + \mathbf{A}\mathbf{K}^H_{\widehat{H}_{d,t}}, \tag{7}$$

where $\mathbf{C}_{t3} \in \mathbb{C}^{\tau_3 \times N_t}$ is the pilot matrix satisfying $\mathbf{C}^H_{t3}\mathbf{C}_{t3} = \mathbf{I}_{N_t}$, and $\mathcal{P}_3$ and τ_3 are the transmission power and training length of the transmitter, respectively. The matrix $\mathbf{A}$ is the AN matrix and its each element is an *i.i.d.* Gaussian random variables with zero mean and variance σ_a^2. $\mathbf{K}_{\widehat{H}_{d,t}} \in \mathbb{C}^{N_t \times (N_t - N_L)}$ is the matrix whose columns consist of the basis of the left null space of $\widehat{\mathbf{H}}_{d,t}$, *i.e.,* $\mathbf{K}^H_{\widehat{H}_{d,t}}\widehat{\mathbf{H}}_{d,t} = \mathbf{0}_{(N_t - N_L)\times N_L}$ and satisfies $\mathbf{K}^H_{\widehat{H}_{d,t}}\mathbf{K}_{\widehat{H}_{d,t}} = \mathbf{I}_{N_t - N_L}$. The received signals at the LR and UR are respectively given by

$$\mathbf{Y}_{L3} = \sqrt{\frac{\mathcal{P}_3\tau_3}{N_t}}\mathbf{C}_{t3}\mathbf{H}_d + \mathbf{A}\mathbf{K}^H_{\widehat{H}_{d,t}}\mathbf{H}_d + \mathbf{W}_3, \tag{8}$$

$$\mathbf{Y}_{U3} = \sqrt{\frac{\mathcal{P}_3\tau_3}{N_t}}\mathbf{C}_{t3}\mathbf{G} + \mathbf{A}\mathbf{K}^H_{\widehat{H}_{d,t}}\mathbf{G} + \mathbf{V}_3, \tag{9}$$

where $\mathbf{W}_3$ and $\mathbf{V}_3$ are the AWGN matrices respectively at the LR and UR and each element of them are assumed to be *i.i.d.* random variables with zero means and noise power σ_w^2 and σ_v^2.

B. Channel Estimation Performance Analysis

In the following, we analyze the channel estimation performances of the UR and LR by assuming that both of them apply the linear minimum mean square error (LMMSE) criterion [8] for channel estimation. To analyze the channel estimation performance of the UR, let us vectorize the received signal of UR (9) as

$$\mathbf{y}_{U3} = \mathrm{vec}(\mathbf{Y}_{U3}) = (\mathbf{I}_{N_U} \otimes \bar{\mathbf{C}}_{t3})\mathbf{g} + (\mathbf{I}_{N_U} \otimes \mathbf{A}\mathbf{K}^H_{\widehat{H}_{d,t}})\mathbf{g} + \mathbf{v}_3 \tag{10}$$

where $\bar{\mathbf{C}}_{t3} = \sqrt{\frac{\mathcal{P}_3\tau_3}{N_t}}\mathbf{C}_{t3}$. Note that the fact $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$ is applied here, and we have $\mathbf{g} = \mathrm{vec}(\mathbf{G})$ and $\mathbf{v}_3 = \mathrm{vec}(\mathbf{V}_3)$. The normalized mean square error (NMSE) of the downlink channel estimate at the UR is given by

$$\mathrm{NMSE}_U = \frac{\mathrm{E}\{(\mathbf{h}_d - \widehat{\mathbf{h}}_d)(\mathbf{h}_d - \widehat{\mathbf{h}}_d)^H\}}{N_t N_L} \tag{11}$$

$$= \frac{\mathrm{Tr}(\sigma_G^2 \mathbf{I}_{N_t N_U} - \mathbf{C}_{g,y_{U3}}\mathbf{C}^{-1}_{y_{U3}}\mathbf{C}^H_{g,y_{U3}})}{N_t N_U} \tag{12}$$

where $\mathbf{C}_{g,y_{U3}} = \sigma_G^2(\mathbf{I}_{N_U} \otimes \bar{\mathbf{C}}^H_{t3})$ is the covariance matrix between $\mathbf{g}$ and $\mathbf{y}_{U3}$, and

$$\mathbf{C}_{y_{U3}} = \mathbf{I}_{N_U} \otimes \left[\sigma_G^2 \bar{\mathbf{C}}_{t3}\bar{\mathbf{C}}^H_{t3} + \left(\sigma_G^2(N_t - N_L)\sigma_a^2 + \sigma_v^2\right)\mathbf{I}_{N_t}\right]$$

is the covariance matrix of $\mathbf{y}_{U3}$. By replacing these covariance matrices into (12), we have

$$\mathrm{NMSE}_U = \left(\frac{1}{\sigma_G^2} + \frac{\mathcal{P}_3\tau_3}{N_t} \frac{1}{\sigma_G^2(N_t - N_L)\sigma_a^2 + \sigma_v^2} \right)^{-1} \tag{13}$$

On the other hand, the NMSE performance of the LR's channel estimation with a generic round-trip pilot matrix cannot be obtained in a closed form. In this work, we aim to provide a design that can be efficiently implemented and its performance can be easily characterized. To this end, we choose the unitary pilot matrix for the round-trip training, i.e., $\mathbf{C}^H_{t0}\mathbf{C}_{t0} = \mathbf{C}_{t0}\mathbf{C}^H_{t0} = \mathbf{I}_{N_t}$. With such a design, an analytical approximation of the LR's channel estimation error performance can be found in (14). The derivation of this approximation is somewhat lengthy and can be found in [9].

IV. Numerical Results and Discussions

In this section, we demonstrate the numerical results on the NMSE performance of the proposed two-way DCE scheme. We consider the MIMO wireless system as that described in Section II with antenna numbers $N_t = 4$, $N_L = 2$ and $N_U = 2$. The channel variances of $\mathbf{H}_u$, $\mathbf{H}_d$ and $\mathbf{G}$ are all set to be one, *i.e.,* $\sigma^2_{H_u} = \sigma^2_{H_d} = \sigma^2_G = 1$. And the noise powers of the

AWGN matrices $\widetilde{\mathbf{W}}$, $\mathbf{W}$ and $\mathbf{V}$ respectively at the transmitter, the LR and the UR are also equal to one, *i.e.,* $\sigma_{\tilde{w}}^2 = \sigma_w^2 = \sigma_v^2 = 1$. Moreover, the training lengths are set to be the antenna number of its corresponding transmitting terminal, *i.e.,* $\tau_0 = 4$, $\tau_2 = 2$ and $\tau_3 = 4$.

In order to facilitate the DCE design, we consider the following optimization problem: To find a set of training powers and AN power $\{\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \sigma_a^2\}$ that minimizes the LR's channel estimation error while confining the UR's channel estimation error above some target threshold. Mathematically, the optimization problem is described as

$$\begin{aligned} \min_{\mathcal{P}_0,\mathcal{P}_1,\mathcal{P}_2,\mathcal{P}_3,\sigma_a^2\geq 0} \quad & \mathrm{NMSE}_L \qquad (17) \\ \text{s.t.} \quad & \mathrm{NMSE}_U \geq \gamma \\ & \mathcal{P}_0\tau_0 + \mathcal{P}_1\tau_0 + \mathcal{P}_2\tau_2 + \mathcal{P}_3\tau_3 \\ & + (N_t - N_L)\sigma_a^2\tau_3 \leq P_{ave}(\tau_0 + \tau_0 + \tau_2 + \tau_3) \end{aligned}$$

where γ is the target minimum threshold on the UR's NMSE. We consider this optimization problem under an average power constraint P_{ave}. With the approximation of the LR's NMSE in (14) and the UR's NMSE in (13), one can numerically find the optimal solution of the power set $\{\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \sigma_a^2\}$. Then Monte-Carlo simulations on the channel estimation and data detection would be used to verify the efficiency of the proposed DCE with the optimized power set. For the numerical results on the channel estimation errors, we also incorporate a lower bound on the NMSE for comparison, which is given as

$$\mathrm{NMSE}_{LB} = \left(\frac{1}{\sigma_{H_d}^2} + \frac{P_{ave}}{N_t\sigma_w^2}\right)^{-1} \qquad (18)$$

which accounts for the minimum achievable NMSE at the LR when the transmitter does not use AN, *i.e.,* $\sigma_a^2 = 0$.

Figure 2 shows the NMSE performance of the LR and UR versus average power constraint P_{ave}. The parameter of γ is set to be 0.03. The NMSE is obtained by averaging over 5000 channel realization in the Monte-Carlo simulation. We see that the NMSE of the UR satisfies the lower limit constraint γ and the NMSE of the LR reduces as the power budget increases and is relatively close to the NMSE lower bound. It is clear that the proposed two-way DCE scheme does discriminate the NMSE performance between the LR and UR. Also, we have confirmed that the approximation given in (14) is very accurate as the values lie on top of the Monti-Carlo simulation results.

Figure 3 shows the symbol error rate (SER) of the data detection at the LR and UR versus the average power constraint P_{ave}. We consider in the data transmission phase the transmitter sends a 4 by 4 complex orthogonal space time block code (OSTBC) with $N_t = 4$, code length being four and containing three 64-QAM source symbols per code block [10]. The data transmission power is set to be P_{ave}. We assume that both the LR and UR employ their channel estimates obtained during the training phase for data detection. The SER is averaging over 100000 channel realization and OSTBCs in the Monte-Carlo simulation. From Fig. 3, we observe that the SER of the LR is relatively close to the one with perfect

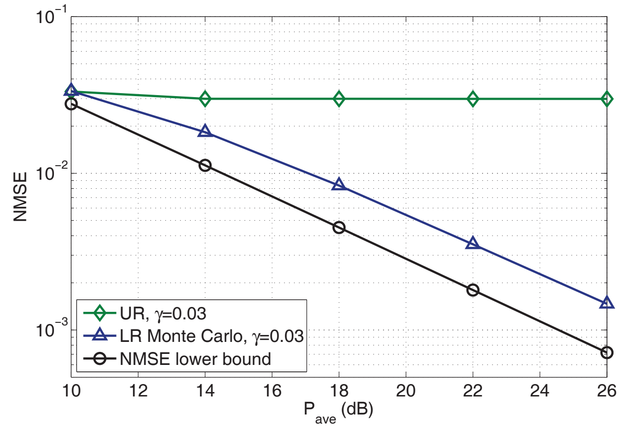

Fig. 2: NMSE performance of the proposed DCE scheme.

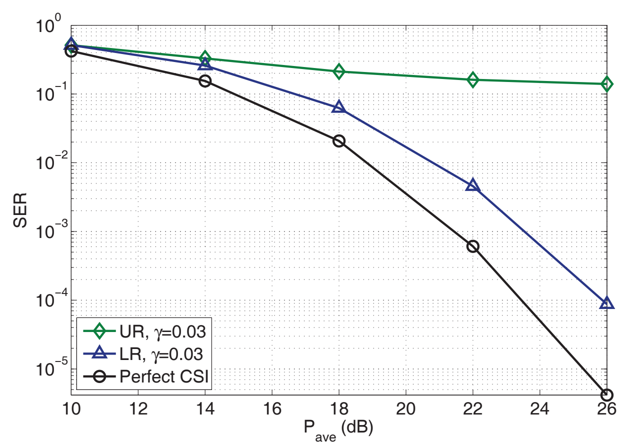

Fig. 3: Symbol error rate of a 64-QAM OSTBC system with the channel estimates obtained by the proposed DCE scheme.

channel state information (CSI). On the other hand, the SER of the UR is confined above 0.1 due to the poor channel estimation performance caused by the proposed DCE scheme. This figure shows that the discrimination between the LR's and UR's data detection performances can be effectively achieved by the proposed two-way DCE scheme.

V. Conclusions

We proposed a new training strategy to achieve DCE in non-reciprocal channels using a two-way training approach. The two-way DCE scheme consists of three training stages, namely the round-trip training, reverse training and forward training. The numerical results verified the efficiency of the proposed DCE scheme in discriminating both the channel estimation performance and the data detection performance between the LR and UR.

APPENDIX

Here we show how the transmitter performs the downlink channel estimation under a given uplink channel estimate. We assume that the transmitter applies the LMMSE criterion for channel estimation. With the help of the reverse training, the transmitter can obtain the uplink channel estimate $\widehat{\mathbf{H}}_u$ which is given by [8]

$$\begin{aligned}\widehat{\mathbf{H}}_u &= \sigma_{H_u}^2 \mathbf{X}_{L2}^H (\sigma_{H_u}^2 \mathbf{X}_{L2} \mathbf{X}_{L2}^H + \sigma_{\tilde{w}}^2 \mathbf{I}_{N_t})^{-1} \mathbf{Y}_{t2} \\ &\triangleq \mathbf{H}_u + \Delta \mathbf{H}_u.\end{aligned}$$

The covariance matrix of the estimation error matrix $\Delta\mathbf{H}_u$ is given by

$$\mathrm{E}\{(\Delta\mathbf{H}_u)^H \Delta\mathbf{H}_u\} = N_L \left(\frac{1}{\sigma_{H_u}^2} + \frac{\mathcal{P}_2 \tau_2}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \mathbf{I}_{N_t} \quad (19)$$

where $\sigma_{\tilde{w}}^2$ is the noise power of the AWGN matrix at the transmitter. With a given $\widehat{\mathbf{H}}_u$ at the transmitter, the echoed signal (3) can be rewritten as

$$\mathbf{Y}_{t1} = \alpha \mathbf{X}_{t0} \mathbf{H}_d \widehat{\mathbf{H}}_u + \alpha (\mathbf{W}_0 \widehat{\mathbf{H}}_u - \mathbf{X}_{t0} \mathbf{H}_d \Delta \mathbf{H}_u - \mathbf{W}_0 \Delta \mathbf{H}_u) + \widetilde{\mathbf{W}}_1 \quad (20)$$

For employing the LMMSE criterion for the downlink channel estimation, it is easier to describe (20) in the vector form as

$$\begin{aligned}\mathbf{y}_{t1} &= \alpha (\widehat{\mathbf{H}}_u^T \otimes \mathbf{X}_{t0}) \mathbf{h}_d + \alpha (\widehat{\mathbf{H}}_u^T \otimes \mathbf{I}_{N_t}) \mathbf{w}_0 \\ &\quad - \alpha (\Delta \mathbf{H}_u^T \otimes \mathbf{X}_{t0}) \mathbf{h}_d - \alpha (\Delta \mathbf{H}_u^T \otimes \mathbf{I}_{N_t}) \mathbf{w}_0 + \tilde{\mathbf{w}}_1\end{aligned}$$

where $\mathbf{y}_{t1} = \mathrm{vec}(\mathbf{Y}_{t1})$, $\mathbf{h}_d = \mathrm{vec}(\mathbf{H}_d)$, $\mathbf{w}_0 = \mathrm{vec}(\mathbf{W}_0)$, and $\tilde{\mathbf{w}}_1 = \mathrm{vec}(\widetilde{\mathbf{W}}_1)$. By the fact that $\widehat{\mathbf{H}}_u$ and $\Delta\mathbf{H}_u$ are uncorrelated due to the orthogonality principle [8] and (19), the LMMSE estimate of dowlink channel downlink channel under a given $\widehat{\mathbf{H}}_u$ is given by

$$\widehat{\mathbf{h}}_{d,t} = \mathbf{C}_{h_d, y_{t1}} \mathbf{C}_{y_{t1}}^{-1} \mathbf{y}_{t1} \quad (21)$$

where $\mathbf{C}_{h_d,y_{t1}} = \mathrm{E}\{\mathbf{h}_d \mathbf{y}_{t1}^H \mid \widehat{\mathbf{H}}_u\} = \alpha \sigma_{H_d}^2 (\widehat{\mathbf{H}}_u^* \otimes \mathbf{X}_{t0}^H)$ is the covariance matrix between $\mathbf{h}_d$ and $\mathbf{y}_{t1}$ and

$$\begin{aligned}\mathbf{C}_{y_{t1}} &= \mathrm{E}\{\mathbf{y}_{t1} \mathbf{y}_{t1}^H \mid \widehat{\mathbf{H}}_u\} \quad (22) \\ &= \alpha^2 \left[\left((\widehat{\mathbf{H}}_u^T \widehat{\mathbf{H}}_u^*) + N_L \left(\frac{1}{\sigma_{H_u}^2} + \frac{\mathcal{P}_2 \tau_2}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \mathbf{I}_{N_t} \right) \right. \\ &\quad \left. \otimes (\sigma_{H_d}^2 \mathbf{X}_{t0} \mathbf{X}_{t0}^H + \sigma_w^2 \mathbf{I}_{\tau_0}) \right] + \sigma_{\tilde{w}}^2 \mathbf{I}_{N_t} \otimes \mathbf{I}_{\tau_0} \quad (23)\end{aligned}$$

is the covariance matrix of $\mathbf{y}_{t1}$. The downlink channel estimate $\widehat{\mathbf{H}}_{d,t}$ can be simply obtained by reshaping its vector form $\widehat{\mathbf{h}}_{d,t}$.

REFERENCES

[1] T.-H. Chang, W.-C. Chiang, Y.-W. P. Hong and C.-Y. Chi, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Process.,* vol. 58, no. 12, pp. 6223-6237, Dec. 2010.

[2] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.,* vol. 7, no. 6, pp. 2180-2189, June 2008.

[3] C. Steger and A. Sabharwal, "Single-input two-way SIMO Channel: diversity-multiplexing tradeoff with two-way training," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4877-4885, Dec. 2008.

[4] X. Zhou, T. A. Lamahewa, P. Sadeghi and S. Durrani, "Two-way training: optimal power allocatin for pilot and data transmission," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 564-569, Feb. 2010.

[5] L. P. Withers, R. M. Taylor and D. M. Warme, "Echo-MIMO: a two-way channel training method for matched cooperative beamforming," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4419-4432, Sep. 2008.

[6] C.-W. Huang, X. Zhou, T.-H. Chang and Y.-W. P. Hong, "Two-Way Training Design for Discriminatory Channel Estimation in Wireless MIMO Systems," *Proc. IEEE International Conference on Communications (ICC)*, Kyoto, Japan, Jun. 2011.

[7] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE. Trans. Inform. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[8] S. M. Kay, *Fundamnetals of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall International, 1993.

[9] C.-W. Huang, T.-H. Chang, X. Zhou and Y.-W. P. Hong, "Two-Way Discriminatory Channel Estimation in Wireless MIMO Systems," submitted to *IEEE Trans. Signal Process.*

[10] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.