# Two-Way Training Design for Discriminatory Channel Estimation in Wireless MIMO systems

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

*Institute of Commun. Eng. & Department of Elect. Eng., National Tsing Hua University, Hsinchu, Taiwan 30013
†UNIK - University Graduate Center, University of Oslo, Kjeller, NO-2027, Norway
Emails: cwhuang@erdos.ee.nthu.edu.tw, xiangyun@unik.no, tsunghui.chang@gmail.com and ywhong@ee.nthu.edu.tw

***Abstract*—This paper examines the use of two-way training in multiple-input multiple-output (MIMO) wireless systems to discriminate the channel estimation (and, thus, data detection) performance between two receivers, namely, a legitimate receiver (LR) and an unauthorized receiver (UR). This work extends upon the discriminatory channel estimation (DCE) proposed in our prior work, where it was previously assumed that training signals can only be sent by the transmitter. The DCE design criterion is to minimize the channel estimation error at the LR while confining the channel estimation error at the UR above a minimum level. In the case of two-way training, training signals can first be transmitted on the reverse link to enable channel estimation at the transmitter and allow the transmitter to insert artificial noise (AN) along with the training signal in the forward link to disrupt the training at the UR, while minimizing the interference on the LR. The optimal power allocation between training and AN signals is devised for systems that are subject to both average and peak power constraints. Numerical results demonstrate the efficacy of the proposed two-way training scheme when used in discriminating the performances between LR and UR.**

## I. Introduction

Secrecy in wireless communications has become an increasingly important issue in recent years due to the broadcast nature of the wireless medium. In the past, these issues have mostly been addressed using cryptography in the application layer. However, recent studies on information-theoretic secrecy provide an alternative to achieve these tasks through coding and modulation in the physical layer. The secrecy capacity, i.e., the rate achievable with vanishing error probability at the legitimate receiver (LR) and vanishing equivocation rate at the unauthorized receiver (UR), has been derived for single-input single-output (SISO) systems in [1] and for multiple-input multiple-output (MIMO) systems in [2]. The results show that secrecy capacity can be increased by enlarging the difference between the effective channel qualities of LR and UR.

While most works on physical-layer secrecy focus on the data transmission design, it has been proposed in [3] a novel idea of achieving performance discrimination in the channel estimation phase. Specifically, in [3], we proposed the so-called discriminatory channel estimation (DCE) scheme, where artificial noise (AN) is superimposed on top of the training signal to disrupt the channel estimation at UR. However, to minimize the interference on LR, the AN signal must be designed based on knowledge of the LR's channel and, thus, the scheme requires LR to feedback its channel estimate during each stage of the process. The main drawback of the DCE scheme in [3] is the need of multiple feedback-and-retraining stages for achieving good performance, which results in significant training overhead and high design complexity.

When the channel is symmetric, e.g., in time-division multiplexing (TDD) systems, the CSI can be obtained at the transmitter by transmitting pilot signals from the receiver. For example, two-way training schemes were studied in [4], [5] to obtain the CSI at both the receiver and the transmitter without the use of feedback. In this paper, we adopt the concept of two-way training into the design to increase the efficiency of the DCE scheme. The proposed two-way DCE scheme uses a reverse training to acquire the CSI at the transmitter and a forward training with AN to achieve different channel estimation qualities at the LR and the UR. Compared to the multi-stage feedback-and-retraining DCE scheme in [3], our proposed scheme drastically decreases the overall training overhead and the design complexity. Furthermore, we propose to optimize the power allocation among the reverse and forward training by solving an optimization problem that aims to minimize the channel estimation error at the LR subject to a lower limit constraint on the channel estimation error at the UR. Our analytical result shows that the problem of finding the optimal reverse training power, forward training power and the AN power can be reformulated as a one-variable problem which can be solved by a simple line search. Numerical results show that the proposed DCE design can effectively discriminate the quality of channel estimation and the data detection performances at the LR and UR.

## II. System Model and Problem Statement

We consider a TDD wireless MIMO system consisting of a transmitter, a legitimate receiver (LR), and an unauthorized receiver (UR). We assume that the transmitter, LR, and UR are equipped with $N_t$, $N_L$ and $N_U$ antennas, respectively. The channels of LR and UR remain constant during one transmission block, which consists of a training phase and a data transmission phase. We assume that space-time block codes (STBC) are used for data transmission and that UR is to passively overhear the information sent from the transmitter

This work was supported in part by the Research Council of Norway through the project 197565/V30 and the National Science Council, Taiwan, under grant s NSC 98-2221-E-007-059-MY2 and NSC 98-2219-E-007-004.

to LR. To prevent UR from extracting the message, we consider the design of a two-way training scheme for DCE that enables LR to perform an accurate estimate of the channel while disrupting the channel estimation performance at UR. Assume that the channel between the transmitter and the LR is reciprocal, which means that the forward channel (i.e., the channel from the transmitter to the LR) and the reverse channel (i.e., the channel from the LR to the transmitter) are symmetric. By denoting the forward channel matrix as $\mathbf{H} \in \mathbb{C}^{N_t \times N_L}$, the reverse channel matrix will be represented as $\mathbf{H}^T \in \mathbb{C}^{N_L \times N_t}$. Utilizing the channel reciprocity, we propose a two-way training scheme for DCE that involves both a reverse and a forward training procedure, as detailed below.

**Reverse training** : First, the LR sends a training signal, denoted by $\mathbf{X}_L \in \mathbb{C}^{T_R \times N_L}$, to enable channel estimation at the transmitter. This allows the transmitter to have knowledge about the LR's channel, and this channel knowledge will be used in the forward training signal design to discriminate between the channel estimation performances of the LR and UR. Specifically, the reverse training signal $\mathbf{X}_L$ is given by

$$\mathbf{X}_L = \sqrt{\frac{P_0 T_R}{N_L}} \mathbf{C}_L, \tag{1}$$

where the pilot matrix $\mathbf{C}_L$ satisfies $\mathbf{C}_L^H \mathbf{C}_L = \mathbf{I}_{N_L}$ (the $N_L$ by $N_L$ identity matrix), and $P_0$ and $T_R$ represent the transmission power and training interval, respectively. The received signal at the transmitter is given by

$$\mathbf{Y}_t = \mathbf{X}_L \mathbf{H}^T + \widetilde{\mathbf{W}}, \tag{2}$$

where each element of $\mathbf{H}$ is assumed to be independent and identically distributed (i.i.d.) random variable with zero mean and variance equal to $\sigma_H^2$, and $\widetilde{\mathbf{W}} \in \mathbb{C}^{T_R \times N_t}$ is the additive white noise matrix with each element having zero mean and variance $\sigma_{\tilde{w}}^2$. We assume that the transmitter employs the linear minimum mean square error (LMMSE) criterion for channel estimation [6]. The channel estimate of $\mathbf{H}$, denoted by $\widehat{\mathbf{H}}$, can be obtained as

$$\begin{aligned} \widehat{\mathbf{H}} &= (\sigma_H^2 \mathbf{X}_L^H (\sigma_H^2 \mathbf{X}_L \mathbf{X}_L^H + \sigma_{\tilde{w}}^2 \mathbf{I}_{T_R})^{-1} \mathbf{Y}_t)^T \\ &\triangleq \mathbf{H} + \Delta\mathbf{H} \end{aligned} \tag{3}$$

where $\Delta\mathbf{H} \in \mathbb{C}^{N_t \times N_L}$ stands for the estimation error matrix. The covariance matrix of $\Delta\mathbf{H}$ can be shown to be [6]

$$\mathrm{E}\{\Delta\mathbf{H}(\Delta\mathbf{H})^H\} = N_L \left( \frac{1}{\sigma_H^2} + \frac{P_0 T_R}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \mathbf{I}_{N_t}. \tag{4}$$

**Forward training** : With the LR's channel estimate $\widehat{\mathbf{H}}$, the transmitter superimposes AN with the training signal in order to degrade the channel estimation performance of the UR. Specifically, by assuming that $N_t > N_L$, the forward training signal is given by

$$\mathbf{X}_t = \sqrt{\frac{P_1 T_F}{N_t}} \mathbf{C}_t + \mathbf{A} \cdot \mathbf{N}_{\hat{H}}^H, \tag{5}$$

where $\mathbf{C}_t \in \mathbb{C}^{T_F \times N_t}$ is the pilot matrix satisfying $\mathbf{C}_t^H \mathbf{C}_t = \mathbf{I}_{N_t}$, $T_F$ is the forward training length, $P_1$ is the transmission power, $\mathbf{A} \in \mathbb{C}^{T_F \times (N_t - N_L)}$ is the AN matrix with each entry being an i.i.d. random variable with zero mean and variance equal to $\sigma_a^2$, and $\mathbf{N}_{\hat{H}} \in \mathbb{C}^{N_t \times (N_t - N_L)}$ is the matrix whose column vectors span the left null space of $\widehat{\mathbf{H}}$, i.e., $\mathbf{N}_{\hat{H}}^H \widehat{\mathbf{H}} = \mathbf{0}$ (the $N_t - N_L$ by $N_L$ zero matrix), and satisfies $\mathbf{N}_{\hat{H}}^H \mathbf{N}_{\hat{H}} = \mathbf{I}_{N_t - N_L}$. As seen in (5), the AN is imposed in the left null space of $\widehat{\mathbf{H}}$ to avoid interfering with the LR. The received signals of LR and UR are respectively given by

$$\mathbf{Y}_L = \sqrt{\frac{P_1 T_F}{N_t}} \mathbf{C}_t \mathbf{H} + \mathbf{A} \cdot \mathbf{N}_{\hat{H}}^H \mathbf{H} + \mathbf{W}, \tag{6}$$

$$\mathbf{Y}_U = \sqrt{\frac{P_1 T_F}{N_t}} \mathbf{C}_t \mathbf{G} + \mathbf{A} \cdot \mathbf{N}_{\hat{H}}^H \mathbf{G} + \mathbf{V}, \tag{7}$$

where $\mathbf{G} \in \mathbb{C}^{N_t \times N_U}$ is the channel matrix from the transmitter to UR, and $\mathbf{W} \in \mathbb{C}^{T_F \times N_L}$ and $\mathbf{V} \in \mathbb{C}^{T_F \times N_U}$ are the additive noise matrices at LR and UR, respectively. Each element of $\mathbf{G}$ is assumed to be i.i.d distributed with zero mean and variance equal to $\sigma_G^2$. Elements of both $\mathbf{W}$ and $\mathbf{V}$ are i.i.d. distributed with zero mean and variances equal to $\sigma_w^2$ and $\sigma_v^2$, respectively.

In the next section, we analyze the channel estimation performances of LR and UR by assuming that both LR and UR employ LMMSE channel estimation. We then propose to judiciously allocate the training powers and the AN power in reverse and forward training, aiming at discriminating between the channel estimation performances of the LR and the UR.

## III. Proposed Two-Way DCE Design

### A. Channel Estimation Performance Analysis

To analyze the channel estimation performance of LR, let us write (6) as

$$\mathbf{Y}_L = \sqrt{\frac{P_1 T_F}{N_t}} \mathbf{C}_t \mathbf{H} - \mathbf{A} \mathbf{N}_{\hat{H}}^H \Delta\mathbf{H} + \mathbf{W} \triangleq \bar{\mathbf{C}} \mathbf{H} + \bar{\mathbf{W}}, \tag{8}$$

where $\bar{\mathbf{C}} \triangleq \sqrt{\frac{P_1 T_F}{N_t}} \mathbf{C}_t$, $\bar{\mathbf{W}} \triangleq -\mathbf{A} \mathbf{N}_{\hat{H}}^H \Delta\mathbf{H} + \mathbf{W}$ and the first equality is due to $\mathbf{N}_{\hat{H}}^H \widehat{\mathbf{H}} = \mathbf{0}$. Denote the channel estimate at LR by $\widehat{\mathbf{H}}_F$. The normalized mean squared error (NMSE) of $\widehat{\mathbf{H}}_F$ under LMMSE criterion can be shown to be [6]

$$\begin{aligned} \mathrm{NMSE}_L &\triangleq \frac{\mathrm{Tr}\left( \mathrm{E}\{(\mathbf{H} - \widehat{\mathbf{H}}_F)(\mathbf{H} - \widehat{\mathbf{H}}_F)^H\} \right)}{N_t N_L} \\ &= \frac{\mathrm{Tr}\left( \left( \mathbf{R}_H^{-1} + \bar{\mathbf{C}}^H \mathbf{R}_{\bar{W}}^{-1} \bar{\mathbf{C}} \right)^{-1} \right)}{N_t N_L}, \end{aligned} \tag{9}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix, $\mathbf{R}_H = N_L \sigma_H^2 \mathbf{I}_{N_t}$ and $\mathbf{R}_{\bar{W}} = \mathrm{E}\{\bar{\mathbf{W}} \bar{\mathbf{W}}^H\}$ is the covariance matrix of $\bar{\mathbf{W}}$. According to the independence between $\mathbf{A}$ and $\mathbf{W}$, the fact of $\mathbf{N}_{\hat{H}}^H \mathbf{N}_{\hat{H}} = \mathbf{I}_{N_t - N_L}$ and (4), it can be shown that

$$\begin{aligned} \mathbf{R}_{\bar{W}} &= \left( \mathrm{E}\{\|\mathbf{N}_{\hat{H}}^H \Delta\mathbf{H}\|^2\} \sigma_a^2 + N_L \sigma_w^2 \right) \mathbf{I}_{T_F} \\ &= N_L \left[ (N_t - N_L) \cdot \left( \frac{1}{\sigma_H^2} + \frac{P_0 T_R}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \sigma_a^2 + \sigma_w^2 \right] \mathbf{I}_{T_F}. \end{aligned} \tag{10}$$

Substituting (10) into (9) yields

$$\text{NMSE}_L = \left( \frac{1}{\sigma_H^2} + \frac{P_1 T_F / N_t}{(N_t - N_L)\left(\frac{1}{\sigma_H^2} + \frac{P_0 T_R}{N_L \sigma_{\tilde{w}}^2}\right)^{-1} \sigma_a^2 + \sigma_w^2} \right)^{-1}. \quad (11)$$

The NMSE performance of the UR can be analyzed in a similar fashion. Specifically, one can show that the NMSE of estimating $\mathbf{G}$ at the UR is given by

$$\text{NMSE}_U = \left( \frac{1}{\sigma_G^2} + \frac{P_1 T_F / N_t}{(N_t - N_L)\sigma_a^2 \sigma_G^2 + \sigma_v^2} \right)^{-1}. \quad (12)$$

### B. Optimal Training Power Allocation

Observing from (11) and (12), the added AN in forward training can affect both the LR and UR's channel estimation performances. To optimize LR's channel estimation performance while preventing the UR from obtaining an accurate estimate of $\mathbf{G}$, we propose to jointly optimize the reverse training power $P_0$, the forward training power $P_1$ and AN power $\sigma_a^2$ by considering the following power allocation problem

$$\begin{aligned} \min_{P_0, P_1, \sigma_a^2 \geq 0} \quad & \text{NMSE}_L \quad (13) \\ \text{s.t.} \quad & \text{NMSE}_U \geq \gamma, \\ & P_0 T_R + (P_1 + (N_t - N_L)\sigma_a^2) T_F \leq \bar{P}_{ave}(T_R + T_F), \\ & P_0 T_R \leq \bar{P}_L T_R, \\ & (P_1 + (N_t - N_L)\sigma_a^2) T_F \leq \bar{P}_t T_F, \end{aligned}$$

where we aim to minimize the LR's NMSE subject to a preset lower limit $\gamma$ on the UR's NMSE, under an average power constraint $\bar{P}_{ave}$. Note that the LR and the transmitter also have their own peak power constraints, i.e., $\bar{P}_L$ and $\bar{P}_t$. To make all constraints effective, we shall focus on the interesting case of

$$\max\{\bar{P}_L T_R,\ \bar{P}_t T_F\} \leq \bar{P}_{ave}(T_R + T_F) \leq \bar{P}_L T_R + \bar{P}_t T_F. \quad (14)$$

Note that for the case of $\bar{P}_{ave}(T_R + T_F) > \bar{P}_L T_R + \bar{P}_t T_F$, the average power constraint becomes redundant and hence, the transmitter and the LR simply transmit with its maximum power. When $\bar{P}_{ave}(T_R + T_F) < \bar{P}_L T_R$ and/or $\bar{P}_{ave}(T_R + T_F) < \bar{P}_t T_F$, one or both individual power constraints become redundant. The solution for this case can be easily obtained by following the derivations for the case of (14)[1].

On the other hand, it should be noted that the preset value $\gamma$ should satisfy [3]

$$\left( \frac{1}{\sigma_G^2} + \frac{\bar{P}_t T_F}{N_t \sigma_v^2} \right)^{-1} \leq \gamma \leq \sigma_G^2, \quad (15)$$

since the left-hand-side term is the minimum achievable NMSE of UR (when the transmitter does not use AN, i.e., $\sigma_a^2 = 0$), and the right-hand-side term stands for the worst NMSE performance of UR, respectively. For ease of latter use, let us define

$$\tilde{\gamma} \triangleq \left( \frac{1}{\gamma} - \frac{1}{\sigma_G^2} \right) N_t \sigma_v^2 \geq 0. \quad (16)$$

Then the condition in (15) reduces to

$$0 \leq \tilde{\gamma} \leq \bar{P}_t T_F. \quad (17)$$

### C. Solving Problem (13) via Line Search

The power allocation problem in (13) is a nonconvex optimization problem involving three variables $(P_0, P_1, \sigma_a^2)$. However, it actually can be solved very efficiently. To see this, let us define $a = P_0 T_R$, $b = P_1 T_F$ and $c = (N_t - N_L)\sigma_a^2$. We can rewrite problem (13) as

$$\begin{aligned} \max_{a,b,c \geq 0} \quad & \frac{(N_L \sigma_w^2 + \sigma_H^2 a) b}{N_L \sigma_w^2 + \sigma_H^2 \cdot a + N_L \sigma_H^2 \frac{\sigma_{\tilde{w}}^2}{\sigma_w^2} \cdot c} \quad (18a) \\ \text{s.t.} \quad & \frac{\sigma_v^2 \cdot b}{\sigma_G^2 \cdot c + \sigma_v^2} \leq \tilde{\gamma}, \quad (18b) \\ & a + b + c \cdot T_F \leq \bar{P}_{ave}(T_R + T_F), \quad (18c) \\ & a \leq \bar{P}_L T_R, \quad (18d) \\ & b + c \cdot T_F \leq \bar{P}_t T_F. \quad (18e) \end{aligned}$$

We show in the Appendix the following proposition for problem (13):

**Proposition 1** *Consider the power allocation problem in* (18) *with both* (14) *and* (17) *satisfied. If*

$$\begin{aligned} \mu \triangleq N_t & \left( \frac{\sigma_v^2 \sigma_{\tilde{w}}^2}{\sigma_G^2 \sigma_w^2} - \frac{\sigma_w^2}{\sigma_H^2} \right) \\ & > \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R + T_F) - \tilde{\gamma}\}, \end{aligned} \quad (19)$$

*the optimal* $(a, b, c)$ *of* (18) *is given by* $a^* = 0$, $b^* = \tilde{\gamma}$ *and* $c^* = 0$ *(i.e., no need of reverse training and no need of AN in forward training). On the other hand, if* $\mu \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R + T_F) - \tilde{\gamma}\}$, *the optimal* $a$ *of* (18) *can be obtained by solving the following one-dimensional problem*

$$\begin{aligned} a^\star = \arg \max_{a \geq 0} \quad & \frac{(N_L \sigma_w^2 + \sigma_H^2 a) b(a)}{N_L \sigma_w^2 + \sigma_H^2 \cdot a + N_L \sigma_H^2 \frac{\sigma_{\tilde{w}}^2}{\sigma_w^2} \cdot c(a)} \quad (20) \\ \text{s.t} \quad & \max\{0, \mu, \bar{P}_{ave}(T_R + T_F) - \bar{P}_t T_F\} \leq a \leq \\ & \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R + T_F) - \tilde{\gamma}\}, \end{aligned}$$

*where*

$$c(a) = \frac{\bar{P}_{ave}(T_R + T_F) - \tilde{\gamma} - a}{T_F + \sigma_G^2 \tilde{\gamma} / \sigma_v^2}, \quad (21)$$

*and*

$$b(a) = \tilde{\gamma} \left( \frac{\sigma_G^2}{\sigma_v^2} \cdot c(a) + 1 \right). \quad (22)$$

*The optimal* $b$ *and* $c$ *are given by* $b^\star = b(a^\star)$ *and* $c^\star = c(a^\star)$.

Proposition 1 implies that the solutions of problem (13) can be efficiently obtained by simple line search over a finite interval, when the condition in (19) is fulfilled; otherwise, one can handily have a simple closed-form solution of $a^* = 0$, $b^* = \tilde{\gamma}$ and $c^* = 0$.

[1]The proposition to be given for the case of (14) also describes the solution for the case of $\bar{P}_{ave}(T_R + T_F) < \bar{P}_L T_R$ and/or $\bar{P}_{ave}(T_R + T_F) < \bar{P}_t T_F$, by changing the condition in (17) to $0 \leq \tilde{\gamma} \leq \min\{\bar{P}_t T_F,\ \bar{P}_{ave}(T_R + T_F)\}$ and setting the redundant individual power constraint(s) to infinity.

## IV. Numerical Results and Discussions

In this section, we present some numerical results on the optimal power allocation and NMSE performance of the proposed DCE scheme. We consider the MIMO wireless system as described in Section II with $N_t = 4$, $N_L = 2$ and $N_U = 2$. The elements of the channel matrices $\mathbf{H}$ and $\mathbf{G}$ are i.i.d. complex Gaussian distributed with zero mean and unit variance ($\sigma_{\widetilde{H}}^2 = \sigma_G^2 = 1$). Each entry of additive white noise matrices $\widetilde{\mathbf{W}}$, $\mathbf{W}$ and $\mathbf{V}$ is also i.i.d. complex Gaussian distributed with zero mean and unit variance, i.e., $\sigma_{\tilde{w}}^2 = \sigma_w^2 = \sigma_v^2 = 1$. Moreover, the training lengths are set to be $T_R = 100$ and $T_F = 100$.[2] Note that the overall training time is larger than $T_R + T_F$ due to the processing time at the transmitter. We incorporate an NMSE lower bound for comparison

$$\text{NMSE}_{\text{LB}} = \left( \frac{1}{\sigma_H^2} + \frac{\min\{\bar{P}_t T_F,\ \bar{P}_{ave}(T_R + T_F)\}}{N_t \sigma_w^2} \right)^{-1} \quad (23)$$

which stands for the minimum achievable NMSE at the LR when $\sigma_a^2 = 0$, i.e., no AN exists.

Figure 1 shows the optimal allocation of the reverse and forward training powers $P_0$, $P_1$ and the AN power $(N_t - N_L)\sigma_a^2$ versus average power constraint $\bar{P}_{ave}$. We compare two different lower limit values $\gamma = 0.1$ and $\gamma = 0.03$, and assign the individual power constraints as $\bar{P}_L = 20$ dB and $\bar{P}_t = 30$ dB. We see from Fig. 1 that it is desirable to allocate more power to the AN and less power to the forward training as $\gamma$ increase from $0.03$ to $0.1$. This is due to the fact that the forward training signal benefits the LR and the UR equally while the AN primarily degrades the UR's estimation performance. In addition, we see that the reverse training power does not change much with $\gamma$, since the reverse training power mainly determines the subspace into which the AN is transmitted, which does not directly influence the NMSE values at the LR and the UR.

Figure 2 shows the NMSE performance of the LR and UR versus average power constraint $\bar{P}_{ave}$. The parameter of $\gamma$, $\bar{P}_L$ and $\bar{P}_t$ are the same as in Fig. 1. We observe that the NMSE of the UR meets the lower limit. Furthermore, the proposed DCE scheme constrains the UR's NMSE well above $\gamma$.

Figure 3 shows the symbol error rate (SER) of the LR and the UR versus the average power constraint $\bar{P}_{ave}$ in the data transmission phase. We consider the scenario that the transmitter sends a $4 \times 4$ complex orthogonal STBC (OSTBC) with $N_t = 4$, the code length is four and containing three 64-QAM source symbols per code block [8]. The data transmission power is set to $\bar{P}_{ave}$. Both the LR and the UR will exploit their channel estimates obtained by the proposed DCE to decode the received symbols. In this Monte-Carlo simulation, the SER is obtained by averaging over $50000$ channel realization and OSTBCs. From Fig. 3, we see that the SER of the LR is quite close to that with perfect CSI while the SER of the UR remains around $0.5$ due to the poor channel

[2]In IEEE 802.11a wireless LAN systems [7], the training sequence length for channel acquisition is around 284 samples.

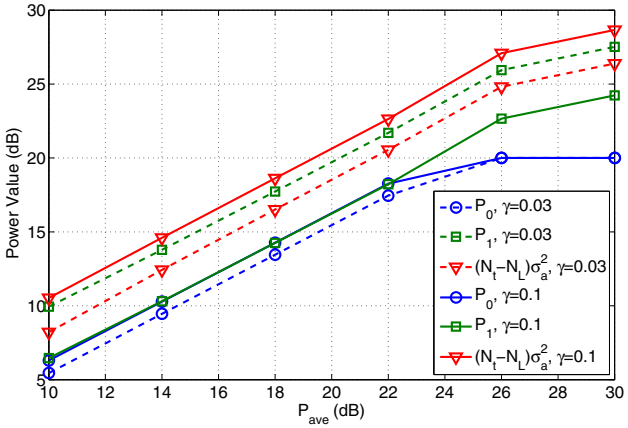


Fig. 1: Power allocation among reverse and forward training powers $P_0$, $P_1$ and AN power $(N_t - N_L)\sigma_a^2$.

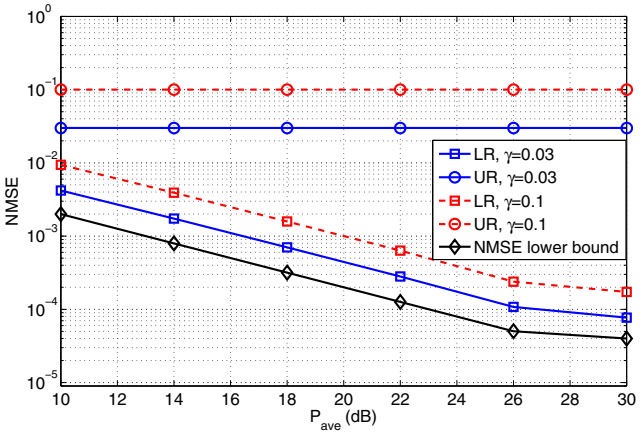


Fig. 2: NMSE performance of the proposed DCE scheme.

estimation performance at the UR. This illustrates that with the proposed two-way training DCE scheme the discrimination of the data detection performances between the LR and the UR can be evidently achieved. It is worthwhile to mention that the feedback-and-retraining DCE scheme proposed in [3] assumes a perfect feedback channel with no power consumption and, thus, it is difficult to have a fair performance comparison between the proposed scheme and that in [3].

## Appendix

The optimization problem can be solved in two steps: (i) find the optimal values of $b$ and $c$ for any given $a$; (ii) find the optimal value of $a$.

**Step i** : we first find the optimal values of $b$ and $c$ as functions of $a$. Note that from (18d) a feasible $a$ must satisfy $a \leq \bar{P}_L T_R$. In the following, we consider two different ranges of $a$.

Case 1 : $\bar{P}_{ave}(T_R + T_F) - \tilde{\gamma} < a \leq \bar{P}_L T_R$, if $\bar{P}_{ave}(T_R + T_F) - \tilde{\gamma} < \bar{P}_L T_R$ holds. Since the objective function in (18a) is monotonically increasing with respect to $b$ but decreasing with respect to $c$, by (17) and the condition of $a > \bar{P}_{ave}(T_R + T_F) - \tilde{\gamma}$, we get $b^*(a) = \bar{P}_{ave}(T_R + T_F) - a$, $c^* = 0$ and

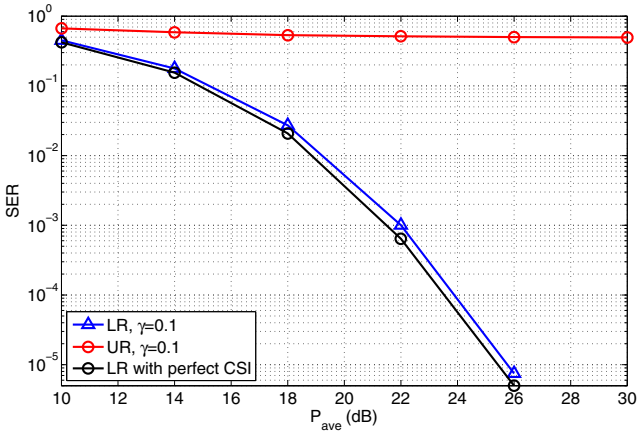


Fig. 3: SER performance of a 64-QAM OSTBC system with the channel estimates obtained by the proposed DCE scheme.

hence the value of (18a) becomes $b^*(a) = \bar{P}_{ave}(T_R+T_F) - a$, which is less than $\tilde{\gamma}$.

Case 2 : $a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$. It can be observed that if the constraint (18b) is inactive we can always decrease $c$ until activating the constraint to obtain a larger objective value. If the condition (18b) is still inactive even when $c = 0$, we can instead lift $b$ to achieve a larger objective value while still satisfying (14), (17) and the condition $a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$. We conclude that constraint (18b) must be active at the optimum. Hence we have

$$b^*(a) = \tilde{\gamma}\left(\frac{\sigma_G^2}{\sigma_v^2}\cdot c^*(a) + 1\right). \tag{24}$$

By substituting (24) into (18), the problem becomes

$$\max_{c\geq 0} \frac{(\sigma_G^2/\sigma_v^2 \cdot c + 1)(N_L\sigma_w^2 + \sigma_H^2 \cdot a)\tilde{\gamma}}{N_L\sigma_H^2\frac{\sigma_{\tilde{w}}^2}{\sigma_w^2}\cdot c + N_L\sigma_w^2 + \sigma_H^2 \cdot a} \tag{25a}$$

$$\text{s.t.} \quad \left(T_F + \frac{\sigma_G^2\tilde{\gamma}}{\sigma_v^2}\right)c + a \leq \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma} \tag{25b}$$

$$\tilde{\gamma}\left(\frac{\sigma_G^2}{\sigma_v^2}c + 1\right) + T_F \cdot c \leq \bar{P}_t T_F. \tag{25c}$$

The range of a in this case is further divided into the following two subranges.

Case 2.1 : when $a < \mu$ and $a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R + T_F) - \tilde{\gamma}\}$ where $\mu \triangleq N_t\left(\frac{\sigma_v^2\sigma_{\tilde{w}}^2}{\sigma_G^2\sigma_w^2} - \frac{\sigma_w^2}{\sigma_H^2}\right)$, the objective function in (25a) is a monotonically decreasing function with respect to $c$. Therefore, the optimal value of $c^*(a)$ is 0 and the corresponding optimal objective value is equal to $\tilde{\gamma}$.

Case 2.2 : as $\mu \leq a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$ the objective function in (25a) is monotonically non-decreasing with respect to $c$. For $\mu \leq a \leq P_{ave}(T_R+T_F) - \bar{P}_t T_F$, constraint (25c) must be active at the optimum with

$$c^* = \frac{\bar{P}_t T_F - \tilde{\gamma}}{T_F + \sigma_G^2\tilde{\gamma}/\sigma_v^2}. \tag{26}$$

Reversely, considering $a \geq \max\{\mu,\ P_{ave}(T_R+T_F) - \bar{P}_t T_F\}$, the constraint (25b) must be active at the optimum with

$$c^*(a) = \frac{P_{ave}(T_R+T_F) - \tilde{\gamma} - a}{T_F + \sigma_G^2\tilde{\gamma}/\sigma_v^2} \tag{27}$$

Moreover, for $a \geq \mu$, the optimal objective value of (25) can be shown to be no less than $\tilde{\gamma}$.

**Step ii** : we now solve for the optimal value of $a$. From the analysis in Step i, a feasible $a$ satisfying $a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$ leads to greater objective value than that of $\bar{P}_{ave}(T_R+T_F) - \tilde{\gamma} < a \leq \bar{P}_L T_R$, thus the optimal value of $a$ must lie in the former condition. For the first case that $\mu > \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$, we can infer $a < \mu$ for all feasible $a$ satisfying $a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$ so that $c^* = 0$ and thus $b^* = \tilde{\gamma}$. Then we get $a^* = 0$ for no need of AN. For the other case of $\mu \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$, from Case 2 we can see that the corresponding objective value for $\max\{0,\ \mu\} \leq a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$ is no less than that for $a < \mu$. If $\mu \leq a \leq \bar{P}_{ave}(T_R+T_F) - \bar{P}_t T_F$ exists, the optimization problem (25) becomes

$$\begin{aligned} \max_{\tilde{a}\geq 0}\ & \frac{(N_L\sigma_w^2 + \sigma_H^2\tilde{a})b^\star}{N_L\sigma_w^2 + \sigma_H^2\cdot\tilde{a} + N_L\sigma_H^2\frac{\sigma_{\tilde{w}}^2}{\sigma_w^2}\cdot c^\star} \\ \text{s.t}\ & \max\{0,\ \mu\} \leq \tilde{a} \leq \bar{P}_{ave}(T_R+T_F) - \bar{P}_t T_F \end{aligned} \tag{28}$$

where $b^\star$ and $c^\star$ are given by (24) and (26) which do not depend on $a$ in this condition. It can be observed that the objective function (28) is monotonically non-decreasing with respect to $\tilde{a}$; thus the optimal value is achieved when $\tilde{a}^* = \bar{P}_{ave}(T_R+T_F) - \bar{P}_t T_F$. However, the corresponding optimal objective value of (28) is the same as the objective value of (20) in this case. Hence, we can have the value of $a^*$ lie in the interval $\max\{0,\ \mu,\ \bar{P}_{ave}(T_R+T_F) - \bar{P}_t T_F\} \leq a \leq \min\{\bar{P}_L T_R,\ \bar{P}_{ave}(T_R+T_F) - \tilde{\gamma}\}$ by solving the optimization problem (20) and the corresponding $c(a)$ and $b(a)$ are given by (27) and (24), respectively.

## REFERENCES


[1] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4687-4698, Oct. 2008.

[2] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas: the MISOME wiretap channel", *IEEE Trans. Inform. Theory,* vol. 56, no. 7, pp. 3088-3104, Jul. 2010.

[3] T.-H. Chang, W.-C. Chiang, Y.-W. Hong and C.-Y. Chi, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," *to appear in IEEE Trans. Signal Process., 2010 (available on IEEE Xplore).*

[4] 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.

[5] 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.

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

[7] IEEE 802.11 Working Group on Broadband Wireless Access, IEEE Standard for Local and Metropolitan Area Networks Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5GHz Band, Sep. 1999.

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