An Improved Two-Way Training for Discriminatory Channel Estimation via Semiblind Approach

Junjie Yang*, Rong Yu*, Xiangyun Zhou †, Yan Zhang ‡
School of Automation, Guangdong University of Technology, China. *
Research School of Engineering, Australian National University, Australia.†
Simula Research Laboratory, Norway. ‡
Email: yangjunjie1985@gmail.com, yurong@gdut.edu.cn, xiangyun.zhou@anu.edu.au, yanzhang@simula.no.

***Abstract*—This paper studies the discriminatory channel estimation (DCE) performance between a legitimate receiver (LR) and an unauthorized receiver (UR) in the multiple-input multiple-output (MIMO) wireless systems. DCE is a recently developed concept that intentionally degrades the channel estimation at the UR so as to minimize the probability of confidential information being eavesdropped by the UR. Usually, the existing DCE scheme is based on the linear minimum mean square error (LMMSE) method with two-way training. In this paper, we propose a new two-way training for DCE based on semiblind approach, *e.g.*, the whitening-rotation (WR)-based channel estimator. To characterize the DCE performance, we derive the closed-form of the normalized mean squared error (NMSE) to the channel estimation at both the LR and the UR. Simulation results show that the proposed two-way training achieves higher performance compared to the two-way training designs in the literature.**

***Keywords*—Physical layer security, Discriminatory channel estimation, semiblind approach.**

I. Introduction

Recently, the secrecy problem has become an increasingly significant topic in the wireless communication. Different from the traditional studies in the higher-layer security technology, e.g., encryption method [1], the physical layer security achieves the communication secrecy by exploiting the physical characteristics of wireless channel. The physical layer security technique has broad application areas, such as vehicular networks [2-4], smart grid[5], M2M networks[6-7], CR networks[8]. From information-theoretical viewpoint, the concept of secrecy capacity [9] lays foundation of physical layer security. The secrecy capacity refers to the maximization of rate attainable at the legitimate receiver (LR) while minimization of rate achievable at the unauthorized receiver (UR), which has been applied into many wireless systems [10,11], e.g., the multiple-input multiple-output (MIMO) systems. In contrast to the theoretical study of physical layer security, the design of training sequence also has attracted lots of concerns like [12-14]. In this paper, we mainly focus on the latter issues of efficiency training design.

Usually, the two-way training schemes [13,14] have been used mainly in the training design for discriminatory channel estimation (DCE) between the LR and the UR. Specifically, the two-way training scheme employs one reverse training pilots to acquire the channel state information (CSI) at the transmitter (TX) and another forward training signals with artificial noise (AN) [15,16] for DCE between the LR and the UR. Without loss of generality, the linear minimum mean-square error (LMMSE) method [17] is a generic channel estimator in the two-way training. To improve the performance of DCE, the study in [13] employs a multi-stage feedback-and-retraining strategy in the signals training. Different from the high complexity of [13], an alternative channel method for the two-way training is considered in this paper, namely, the whitening-rotation (WR)[18-20]-based channel estimator. The proposed scheme with new channel estimator adopts a sequence of Gaussian randomized signals within the reverse training phase, and employs a given training pilots with AN during the forward training phase.

In summary, the contributions of this paper can be given as follows:

- The proposed semiblind-based two-way training has a better channel estimation performance than the LMMSE-based two-way training. A closed-form of the normalized mean squared error (NMSE) criterion is derived to evaluate the channel performance between the LR and the UR, respectively. Based on the analytical result, we show that the WR-based semiblind estimator has a substantial DCE performance than the two-way training in the literature.
- Different from the existing power allocation methods, the power allocation problem is solved separately to the terminals at LR and TX. Specifically, the optimization problem is divided into two subproblems: 1) the reverse training power is fixed as a constant value; and 2) the power allocation to the forward training signals involving the pure training pilots and AN can be found by a simple concave optimization with one dimension line search.

The remainder of this paper is organized as follows. First, the system model is offered and the improved two-way training scheme is discussed in section II. Next the performance analysis of two-way training and power allocation problem are discussed in section III. The numerical results are presented in section IV. Finally, the conclusion are drawn in Section V.

II. System model and proposed Discriminatory channel estimation

A. System model and problem description

Fig.1 shows a wireless MIMO system consisting of a transmitter (TX), a legitimate receiver (LR) and an unauthorized receiver (UR). The TX, the LR and the UR have N_T, N_L, and N_U antennas ($N_T > N_L$ and $N_T \geq N_U$), respectively. Matrix symbol H represents the legitimate channel between the TX and the LR while matrix G refers to the wiretap channel between the TX and the UR. H and G both are the Rayleigh flat fading channels [21] in which the elements of channels are

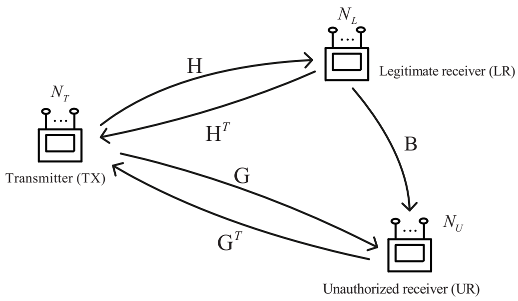


Fig. 1. A wireless MIMO system including a multi-antennas transmitter (TX), a multi-antennas legitimate receiver (LR) and a multi-antennas unauthorized receiver (UR).

assumed to be independent and identifiably distributed random variable with zero mean and variance of $\sigma_{\rm H}^2$, $\sigma_{\rm G}^2$, respectively. Besides, the channel is assumed to be symmetric, *e.g.*, H denotes the channel from the TX to the LR, and $\rm H^T$ refers to the channel from the LR to the TX.

The task of the TX is enabling the LR to interpret the legitimate channel H but also degrading the wiretap channel estimation at the UR through the training sequence design. For enhancing the DCE performance in the noisy case, we address an improved version of two-way training scheme using the WR-based semiblind channel estimator. As the alternative channel estimator, we give a brief introduction of the WR-based method. For example, the uplink channel $\rm H^{\it T}$ can be decomposed as follows

$$\rm H^{\it T} = WQ^{\it H}, \tag{1}$$

where $\rm W \in \mathbb{C}^{N_T \times N_L}$ is a whitening matrix and $\rm Q \in \mathbb{C}^{N_L \times N_L}$ is an unitary rotation matrix, *i.e.*, $\rm Q^{\it H}Q = QQ^{\it H} = I_{N_L}$. In particular, the whitening matrix W can be obtained by using a subspace-based method to the autocorrelation matrix of the received signals; and the rotation matrix Q can be estimated by a constrained maximum likelihood (ML)-based method.

In general, the data training phases include two phases, that is reverse training and forward training. First, the LR transmits a sequence of stochastic signals, only known by itself, in the reverse training phase, facilitating CSI acquisition at the TX; and then the TX broadcasts a new pilot signals inserted with AN in the forward training phase, enabling the channel estimation at the LR whilst disrupting the channel estimation at the UR.

B. Reverse training phase

In the reverse training phase, the LR sends a randomized sequence $\rm S_0 \in \mathbb{C}^{N_L \times T_0}$ back to the TX, where each column of $\rm S_0$ are independent identically distributed (*i.i.d.*) Gaussian variables with zero mean and power variance P_0. The received signals have the following expression,

$$\rm X_0 = H^{\it T}S_0 + E_0, \tag{2}$$

where $\rm X_0 \in \mathbb{C}^{N_T \times T_0}$ and $\rm E_0 \in \mathbb{C}^{N_T \times T_0}$ is the additive Gaussian noises with zero mean and variance σ_0^2.

In general, the whitening matrix can be estimated by performing singular value decomposition (SVD) [22] to the autocorrelation matrix of $\rm X_0$ as follows

$$\rm R_{X_0} \triangleq \frac{X_0 X_0^{\it H}}{\it T_0}. \tag{3}$$

Notice that Eq.(3) can be rewritten as

$$\rm R_{X_0} = \it P_0 \rm H^{\it T}H^* + \Delta R_{X_0}, \tag{4}$$

where ΔR_{X_0} refers to the perturbation error due to the finite length of $\rm S_0$. According to the results of [20], two specific training sequences $\rm S_0^{pos}$ and $\rm S_0^{neg}$ are added into the training sequences to cancellate the perturbation error of $\rm \Delta R_{X_0}$.

By denoting the corresponding received data at the TX as the $\rm X_0^{pos}$ and $\rm X_0^{neg}$, error $\rm \Delta R_{X_0}$ can be calculated by

$$\rm \Delta R_{X_0} = \eta[(R_{X_0^{pos}} - R_{X_0^{neg}}) - \frac{(\it T_0^{\rm pos} - T_0^{\rm neg})}{\it T_0}\sigma_0^2 I_{\it N_T}], \tag{5}$$

where η is the scaling factor. Therefore, the corrected autocorrelation matrix has the form of

$$\rm R_{\bar{X}_0} = R_{X_0} - \Delta R_{X_0}. \tag{6}$$

Using the result of (6), the whitening matrix can be estimated as

$$\rm \hat{W}_0 = \frac{1}{\sqrt{\it P_0}} U_{\bar{X}_0} \Sigma_{\bar{X}_0}^{\frac{1}{2}} \tag{7}$$

by performing an SVD to the $\rm R_{\bar{X}_0}$, where

$$\rm R_{\bar{X}_0} = U_{\bar{X}_0} \Sigma_{\bar{X}_0} V_{\bar{X}_0}.$$

Due to the interference of noises, we define the perturbation error of matrix $\rm \hat{W}_0$ as $\rm \Delta W_0 \triangleq \hat{W}_0 - W$. Using the result of [19], the error of estimate whitening matrix can be deduced as

$$\rm \Delta W_0 \approx \frac{1}{\it P_0} \Delta R_{S_0,E_0}^{\it H} Q, \tag{8}$$

where $\rm \Delta R_{S_0,E_0}$ is the cross correlation matrix between $\rm S_0$ and $\rm E_0$.

C. Forward training phase

In the forward training phase, the AN-aided training sequence is introduced to enable the LR to interpret the channel information while degrading the channel performance at the UR. Specifically, we define a pure training signal as

$$\rm \tilde{S}_1 = \sqrt{\frac{\it P_1}{N_T} \it T_1} C, \tag{9}$$

where P_1 is the pure training data power, $\rm C \in \mathbb{C}^{N_T \times T_1}$ represents the training data matrix, e.g., satisfying orthogonal condition $\rm CC^{\it H} = I_{\it N_T \times N_T}$. To degrade the UR's channel estimate performance, we offer the complete training signals with the form of

$$\rm S_1 \triangleq \tilde{S}_1 + N_{\hat{W}_0} A, \tag{10}$$

where $\rm N_{\hat{W}_0} \in \mathbb{C}^{N_L \times (N_T - N_L)}$ is the orthogonal complement space matrix of $\rm \hat{W}_0$ satisfying that $\rm N_{\hat{W}_0}^{\it H} N_{\hat{W}_0} = I_{\it N_T - N_L}$, and $\rm A \in \mathbb{C}^{(\it N_T - N_L) \times T_1}$ is AN matrix with each component being i.i.d. zero-mean, complex Gaussian random variables with variance σ_a^2.

Using the AN-aided training sequence $\rm S_1$, the received signals at the LR and the UR have the following form, respectively

$$\mathrm{X}_1 = \mathrm{H}\mathrm{S}_1 + \mathrm{E}_1, \tag{11}$$

$$\mathrm{Y}_1 = \mathrm{G}\mathrm{S}_1 + \mathrm{F}_1, \tag{12}$$

where $\mathrm{E}_1 \in \mathbb{C}^{N_L \times T_1}$ and $\mathrm{F}_1 \in \mathbb{C}^{N_U \times T_1}$ are the Gaussian noises with the same variance σ_0^2. Both the LR and the UR make use of the acquired signals for the respective channel estimation, but the performance of the UR will be degraded due to AN.

- **The channel estimation at the LR**

For simplicity, the received signal matrix of Eq.(11) can be rewritten as,

$$\mathrm{LR}: \mathrm{X}_1 = \mathrm{H}\tilde{\mathrm{S}}_1 + \tilde{\mathrm{E}}_1, \tag{13}$$

where $\tilde{\mathrm{E}}_1 \triangleq \mathrm{H}\mathrm{N}_{\hat{\mathrm{W}}_0}\mathrm{A} + \mathrm{E}_1$. Then, we imply the semiblind-based channel method for the channel estimation at the LR. First, the whitening matrix can be obtained as $\hat{\mathrm{W}}_1 = \mathrm{V}^*_{\hat{\mathrm{X}}_\mathrm{W}}\Sigma^T_{\hat{\mathrm{X}}_\mathrm{W}}$ by performing an SVD to the matrix

$$\hat{\mathrm{X}}_\mathrm{W} \triangleq \frac{\mathrm{X}_1\tilde{\mathrm{S}}_1^H}{\frac{P_1}{N_\mathrm{T}}T_1}, \tag{14}$$

where $\hat{\mathrm{X}}_\mathrm{W} = \mathrm{U}_{\hat{\mathrm{X}}_\mathrm{W}}\Sigma_{\hat{\mathrm{X}}_\mathrm{W}}\mathrm{V}^H_{\hat{\mathrm{X}}_\mathrm{W}}$. Next, the rotation matrix can be obtained by solving the following optimization problem under the perturbation-free case,

$$\begin{aligned} LR: \quad & min f(\mathrm{Q}) = \sum_{i=1}^{N_L} \| \mathrm{X}_1(i) - \sum_{j=1}^{N_T} \hat{\sigma}_j q_{ij}^* \hat{\mathrm{S}}_1(j) \|_F^2 \\ \text{s.t.} \quad & \mathrm{Q}\mathrm{Q}^H = \mathrm{I}_{N_L}, \end{aligned} \tag{15}$$

where $\hat{\mathrm{S}}_1 = \mathrm{V}^H_{\mathrm{X}_\mathrm{W}}\tilde{\mathrm{S}}_1$, $\mathrm{X}_1(i)$ is the ith column of X_1 and $\hat{\mathrm{S}}_1(j)$ is the jth column of $\hat{\mathrm{S}}_1$, respectively. By using the Lagrange method, the estimate rotation matrix in terms of $\hat{\mathrm{Q}}_1$ can be obtained $\hat{\mathrm{Q}}_1 = \mathrm{U}_{\hat{\mathrm{X}}_\mathrm{Q}}\mathrm{V}^H_{\hat{\mathrm{X}}_\mathrm{Q}}$ by performing an SVD to the matrix

$$\hat{\mathrm{X}}_\mathrm{Q} \triangleq \frac{\mathrm{X}_1^*\tilde{\mathrm{S}}_1^T\hat{\mathrm{W}}_1}{\frac{P_1}{N_T}T_1}, \tag{16}$$

where $\hat{\mathrm{X}}_\mathrm{Q} = \mathrm{U}_{\hat{\mathrm{X}}_\mathrm{Q}}\Sigma_{\hat{\mathrm{X}}_\mathrm{Q}}\mathrm{V}^H_{\hat{\mathrm{X}}_\mathrm{Q}}$. According to (2), the estimated channel has the expression

$$\hat{\mathrm{H}}_1 = \hat{\mathrm{Q}}_1^*\hat{\mathrm{W}}_1^T. \tag{17}$$

- **The channel estimation at the UR**

Similar to Eq.(2), channel G can be decomposed as

$$\mathrm{G} = \mathrm{M}\mathrm{R}^H, \tag{18}$$

where $\mathrm{M} \in \mathbb{C}^{N_U \times N_T}$ is the whitening matrix, and $\mathrm{R} \in \mathbb{C}^{N_T \times N_T}$ is the rotation matrix. In particular, the received signal matrix of Eq.(12) can be rewritten as,

$$\mathrm{UR}: \mathrm{Y}_1 = \mathrm{G}\tilde{\mathrm{S}}_1 + \tilde{\mathrm{F}}_1, \tag{19}$$

where $\tilde{\mathrm{F}}_1 \triangleq \mathrm{G}\mathrm{N}_{\hat{\mathrm{W}}_0}\mathrm{A} + \mathrm{F}_1$. The whitening matrix can be obtained by

$$\hat{\mathrm{M}} = \mathrm{U}_{\hat{\mathrm{Y}}_\mathrm{M}}\Sigma_{\hat{\mathrm{Y}}_\mathrm{M}} \tag{20}$$

by performing an SVD to the matrix

$$\hat{\mathrm{Y}}_\mathrm{M} \triangleq \frac{\mathrm{Y}_1\tilde{\mathrm{S}}_1^H}{\frac{P_1}{N_T}T_1}, \tag{21}$$

where $\hat{\mathrm{Y}}_\mathrm{M} = \mathrm{U}_{\hat{\mathrm{Y}}_\mathrm{M}}\Sigma_{\hat{\mathrm{Y}}_\mathrm{M}}\mathrm{V}^H_{\hat{\mathrm{Y}}_\mathrm{M}}$. Next, the estimated rotation matrix can be calculated with the training-based method [19],

$$\hat{\mathrm{R}} = \mathrm{V}_{\hat{\mathrm{Y}}_\mathrm{R}}\mathrm{U}^H_{\hat{\mathrm{Y}}_\mathrm{R}} \tag{22}$$

by performing an SVD to the matrix

$$\hat{\mathrm{Y}}_\mathrm{R} \triangleq \frac{\hat{\mathrm{M}}^H\mathrm{Y}_1\tilde{\mathrm{S}}_1^H}{\frac{P_1}{N_T}T_1}, \tag{23}$$

where $\hat{\mathrm{Y}}_\mathrm{R} = \mathrm{U}_{\hat{\mathrm{Y}}_\mathrm{R}}\Sigma_{\hat{\mathrm{Y}}_\mathrm{R}}\mathrm{V}^H_{\hat{\mathrm{Y}}_\mathrm{R}}$. In consequence, the estimated wiretap channel matrix G can be calculated as

$$\hat{\mathrm{G}} = \hat{\mathrm{M}}\hat{\mathrm{R}}^H. \tag{24}$$

III. Performance Analysis of improved two-way Training

A. The Channel Estimation Performance at the LR

We define the perturbation errors of $\hat{\mathrm{W}}_1$ and $\hat{\mathrm{Q}}_1$ as, namely $\Delta\mathrm{W}_1 \triangleq \hat{\mathrm{W}}_1 - \mathrm{W}$ and $\Delta\mathrm{Q}_1 \triangleq \hat{\mathrm{Q}}_1 - \mathrm{Q}$. In the following, the estimation error of channel H at the LR is given by

$$\begin{aligned} \Delta\mathrm{H}_1 &\triangleq \hat{\mathrm{H}}_1 - \mathrm{H} \\ &= \hat{\mathrm{Q}}^*\hat{\mathrm{W}}_1^T - \mathrm{Q}^*\mathrm{W}_1^T \approx \mathrm{Q}^*\Delta\mathrm{W}_1^T + \Delta\mathrm{Q}_1^*\mathrm{W}^T. \end{aligned} \tag{25}$$

First, we derive the close-form expression of error $\Delta\mathrm{W}_1$. Similar to Eq.(8), we have

$$\Delta\mathrm{W}_1 = \frac{N_T}{P_1}\mathrm{Q}^T\Delta\mathrm{R}^*_{\tilde{\mathrm{S}}_1,\tilde{\mathrm{E}}_1}, \tag{26}$$

where

$$\Delta\mathrm{R}_{\tilde{\mathrm{S}}_1,\tilde{\mathrm{E}}_1} = \Delta\mathrm{R}_{\tilde{\mathrm{S}}_1,\mathrm{E}_1} + \Delta\mathrm{R}_{\tilde{\mathrm{S}}_1,\mathrm{A}}\mathrm{N}^H_{\hat{\mathrm{W}}_0}\mathrm{H}^H.$$

Notice that $\mathrm{N}^T_{\hat{\mathrm{W}}_0}\hat{\mathrm{W}}_0 = 0$, Eq.(26) can be rewritten as

$$\Delta\mathrm{W}_1 = \frac{N_T}{P_1}\mathrm{Q}^T(\Delta\mathrm{R}^*_{\tilde{\mathrm{S}}_1,\mathrm{E}_1} - \frac{1}{P_0}\Delta\mathrm{R}^*_{\tilde{\mathrm{S}}_1,\mathrm{A}}\mathrm{N}^T_{\hat{\mathrm{W}}_0}\Delta\mathrm{R}^H_{\mathrm{S}_0,\mathrm{E}_0}).$$

Next, we deduce the close-form expression of error $\Delta\mathrm{Q}_1$. Using the result of [20], we have

$$\Delta\mathrm{Q}_1 \approx \mathrm{Q}(\Gamma_\mathrm{Q} \circ \Pi_\mathrm{Q})^H, \tag{27}$$

where $\circ$ represents the Hadamard product,

$$\Gamma_\mathrm{Q} = \begin{bmatrix} \frac{1}{2\sigma_1^2}, & \cdots & , \frac{1}{\sigma_1^2+\sigma_{N_L}^2} \\ \frac{1}{\sigma_2^2+\sigma_1^2}, & \cdots & , \frac{1}{\sigma_2^2+\sigma_{N_L}^2} \\ & \cdots & \\ \frac{1}{\sigma_{N_L}^2+\sigma_1^2}, & \cdots & , \frac{1}{2\sigma_{N_L}^2} \end{bmatrix} \tag{28}$$

and

$$\Pi_\mathrm{Q} = \Delta\mathrm{X}_\mathrm{Q}^H\mathrm{Q} - \mathrm{Q}^H\Delta\mathrm{X}_\mathrm{Q}. \tag{29}$$

To derive the perturbation error of $\Delta\mathrm{X}_\mathrm{Q}$, we define that $\hat{\mathrm{X}}_\mathrm{Q} \triangleq \mathrm{X}_\mathrm{Q} + \Delta\mathrm{X}_\mathrm{Q}$. Eq.(16) can be modified as

$$\hat{\mathrm{X}}_\mathrm{Q} = \mathrm{H}^*\mathrm{W} + \mathrm{H}^*\Delta\mathrm{W}_1 + \frac{N_T}{P_1}\Delta\mathrm{R}^T_{\tilde{\mathrm{S}}_1,\tilde{\mathrm{E}}_1}\mathrm{W}, \tag{30}$$

Therefore, we have

$$\Delta \mathrm{X_Q} = \mathrm{H}^* \Delta \mathrm{W}_1 + \frac{N_T}{P_1} \Delta \mathrm{R}^T_{\tilde{\mathrm{S}}_1, \tilde{\mathrm{E}}_1} \mathrm{W}. \tag{31}$$

Substituting (31) into (29), we have $\Pi_\mathrm{Q} = 0$. As a result, Eq.(27) can be summarized as $\Delta \mathrm{Q}_1 = 0$, and the perturbation error of channel H can be derived by

$$\Delta \mathrm{H}_1 = \frac{P_1}{N_T} (\Delta \mathrm{R}^*_{\tilde{\mathrm{S}}_1, \mathrm{E}_1} - \frac{1}{P_0} \Delta \mathrm{R}^*_{\tilde{\mathrm{S}}_1, \mathrm{A}} \mathrm{N}^T_{\hat{\mathrm{W}}_0} \Delta \mathrm{R}^H_{\mathrm{S}_0, \mathrm{E}_0}). \tag{32}$$

In consequence, the associate NMSE criterion of $\hat{\mathrm{H}}_1$ can be shown as

$$\begin{aligned} NMSE_L &\triangleq \frac{Tr(\mathrm{E}\{\Delta \mathrm{H}_1 \Delta \mathrm{H}_1^H\})}{N_L N_T} \\ &= \frac{N_T \sigma_0^2}{P_1 T_1} + \frac{(N_T - N_L)\sigma_a^2}{P_0 T_0} \frac{N_T \sigma_0^2}{P_1 T_1}. \end{aligned} \tag{33}$$

B. The Channel Estimation Performance at the UR

In the same manner, we give the definition of perturbation errors of $\hat{\mathrm{M}}_1$ and $\hat{\mathrm{R}}_1$, namely, $\Delta \mathrm{M}_1 \triangleq \hat{\mathrm{M}}_1 - \mathrm{M}$ and $\Delta \mathrm{R}_1 \triangleq \hat{\mathrm{R}}_1 - \mathrm{R}$. In the following, the estimation error of channel G can be written as

$$\Delta \mathrm{G} \triangleq \hat{\mathrm{G}} - \mathrm{G} = \hat{\mathrm{M}}\hat{\mathrm{R}}^H - \mathrm{M}\mathrm{R}^H \approx \mathrm{M}\Delta \mathrm{R}^H + \Delta \mathrm{M}\mathrm{R}^H. \tag{34}$$

First, we derive the closed-form expression of error matrix $\Delta\mathrm{M}$. Similar to Eq. (8), $\Delta\mathrm{M}$ has the following expression

$$\Delta \mathrm{M} = \frac{N_T}{P_1} \Delta \mathrm{R}^H_{\tilde{\mathrm{S}}_1, \tilde{\mathrm{F}}_1} \mathrm{R}, \tag{35}$$

where $\Delta \mathrm{R}_{\tilde{\mathrm{S}}_1, \tilde{\mathrm{F}}_1} \triangleq \Delta \mathrm{R}_{\tilde{\mathrm{S}}_1, \mathrm{A}} \mathrm{N}^H_{\hat{\mathrm{W}}_0} \mathrm{G}^H + \Delta \mathrm{R}_{\tilde{\mathrm{S}}_1, \mathrm{F}_1}$. Next, we deduce the closed-form expression of $\Delta\mathrm{R}$. Similar to Eq.(27), we have

$$\Delta \mathrm{R} \approx \mathrm{R}(\Gamma_\mathrm{R} \circ \Pi_\mathrm{R}), \tag{36}$$

where the structure of Γ_R is similar to the Γ_Q and

$$\Pi_\mathrm{R} = \mathrm{R}^H \Delta \mathrm{Y}^H_\mathrm{R} - \Delta \mathrm{Y}_\mathrm{R} \mathrm{R}. \tag{37}$$

To derive the perturbation error of $\Delta \mathrm{Y_R}$, we define $\hat{\mathrm{Y}}_\mathrm{R} \triangleq \mathrm{Y_R} + \Delta \mathrm{Y_R}$. Besides, Eq. (23) can be approximately by

$$\hat{\mathrm{Y}}_\mathrm{R} \approx \mathrm{M}^H \mathrm{G} + \Delta \mathrm{M}^H \mathrm{G} + \frac{N_\mathrm{T}}{P_1} \mathrm{M}^H \Delta \mathrm{R}^H_{\tilde{\mathrm{S}}_1, \tilde{\mathrm{F}}_1}, \tag{38}$$

Therefore,

$$\Delta \mathrm{Y_R} = \Delta \mathrm{M}^H \mathrm{G} + \frac{N_T}{P_1} \mathrm{M}^H \Delta \mathrm{R}^H_{\tilde{\mathrm{S}}_1, \tilde{\mathrm{F}}_1}. \tag{39}$$

Substituting (39) into (37), we have $\Pi_\mathrm{R} = 0$. As a result, Eq.(36) can be summarized as $\Delta \mathrm{R} = 0$. Then the perturbation error matrix of $\hat{\mathrm{G}}$ can be given by

$$\Delta \mathrm{G} = \frac{N_T}{P_1} (\mathrm{G}\mathrm{N}_{\hat{\mathrm{W}}_0} \Delta \mathrm{R}_{\tilde{\mathrm{S}}_1, \mathrm{A}} + \Delta \mathrm{R}_{\tilde{\mathrm{S}}_1, \mathrm{F}_1}). \tag{40}$$

In consequence, the associate NMSE criterion of $\hat{\mathrm{G}}$ can be shown as

$$\begin{aligned} NMSE_U &\triangleq \frac{Tr\{\mathrm{E}(\Delta \mathrm{G} \Delta \mathrm{G}^H)\}}{N_\mathrm{T} N_\mathrm{U}} \\ &= \frac{N_T \sigma_0^2 + N_T (N_T - N_L) \sigma_a^2 \sigma_G^2}{P_1 T_1}. \end{aligned} \tag{41}$$

C. Optimal Training Power Allocation

We divide the power allocation problem as two subproblems: 1) the power of P_0 can be set as fixed value such as average power P_{ave}; 2) and the power allocation between power P_1 and σ_a^2 can be found by the following formulation

$$\begin{aligned} \min_{P_1 > 0, \sigma_a^2 \geqslant 0} \quad & NMSE_L \\ s.t. \quad & NMSE_U \geq \gamma, \\ & P_1 + (N_T - N_L)\sigma_a^2 \leq P_{ave}. \end{aligned} \tag{42}$$

where $\gamma > 0$ is the preassigned lower limit on the UR's achievable NMSE, and the second constraint of (42) is the total energy constraint to the training signal. Let us define variables as: $x = \frac{P_1 T_1}{N_T}, y = (N_T - N_L)\sigma_a^2$, constant parameter $c = P_0 T_0$. Then the problem of (42) can be reformulated as

$$\begin{aligned} \min_{x > 0, y \geq 0} \quad & \frac{\sigma_0^2}{x} + \frac{y \sigma_0^2}{x} (1 + \frac{N_U \sigma_G^2}{N_T T_0}) \\ s.t. \quad & \frac{\sigma_0^2}{x} + \frac{y \sigma_G^2}{x} \geq \gamma, \\ & \frac{x N_T}{T_1} + y \leq P_{ave}. \end{aligned} \tag{43}$$

Model (43) is a convex optimization problem which involves two variables (x, y), and we give an equivalently form of proposition 1 to solve this optimum problem.

- ***Proposition 1*** *Let $\{x^*, y^*\}$ be the optimal solution to the convex optimization problem in (43) with the constraint $\frac{N_T \sigma_0^2}{P_{ave} T_1} \leq \gamma \leq (N_T - N_L) P_{ave}$. The optimal value of x can be solved by the following one-dimensional problem as follows.*

$$\begin{aligned} \min_x \quad & \frac{\sigma_0^2}{x} + \frac{y(x) \sigma_0^2}{x} (1 + \frac{N_U \sigma_G^2}{N_T T_0}) \\ s.t. \quad & \frac{\sigma_0^2}{\gamma} \leq x \leq \frac{(\sigma_G^2 P_{ave} + \sigma_0^2) T_1}{\gamma T_1 + N_T \sigma_G^2}, \end{aligned} \tag{44}$$

where

$$y(x) = \frac{x\gamma - \sigma_0^2}{\sigma_G^2}. \tag{45}$$

The associated value of y^* is given by $y(x^*)$.

The proof of proposition 1 can be found in the Appendix A. To obtain the optimal result, the linear search like [23] can be employed to solve proposition 1.

IV. SIMULATION RESULTS

In this section, simulation results are provided to demonstrate the validity of proposed scheme comparing with the existing scheme like [13]. For simplicity, the proposed DCE scheme is named as the SPF-WR-based DCE scheme (SPF-WR/S) and the schemes in [13] is called as the LMMSE-based DCE scheme (LMMSE/S). We consider a MIMO system as described in Section II with $N_T = 4, N_L = 2, N_U = 2$, respectively. Channel matrices H and G are randomly generated as *i.i.d* complex Gaussian matrices with zero mean and variance $\sigma_\mathrm{H}^2 = \sigma_\mathrm{G}^2 = 1$. And the additive noise matrices $\mathrm{E}_0, \mathrm{E}_1, \mathrm{F}_1$ are also *i.i.d* complex Gaussian distribution with zero means and unit variance $\sigma_0^2 = 0.01$. We set the maximum

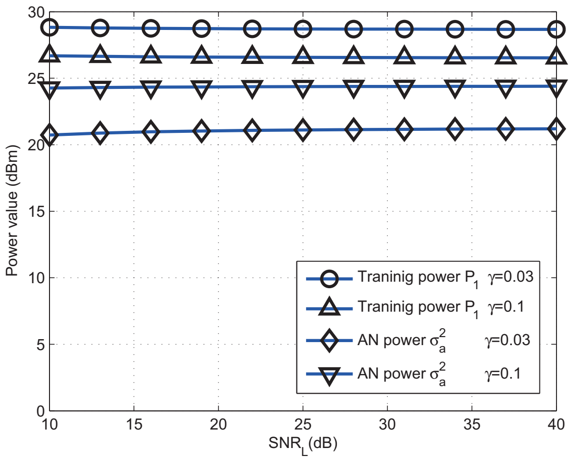


Fig. 2. Power allocation of forward training powers P_1 and σ_a^2.

transmission power as $30dBm$, *i.e.*, $P_{ave} = 1$. Moreover, the overall training lengths are fixed as $T = 280$, in which $T_0 = 136, T_{pos} + T_{neg} = 4$ and $T_1 = 140$. To evacuate the DCE performance of each method, we incorporate the criterion of signal-to-noise ratios (SNR) as

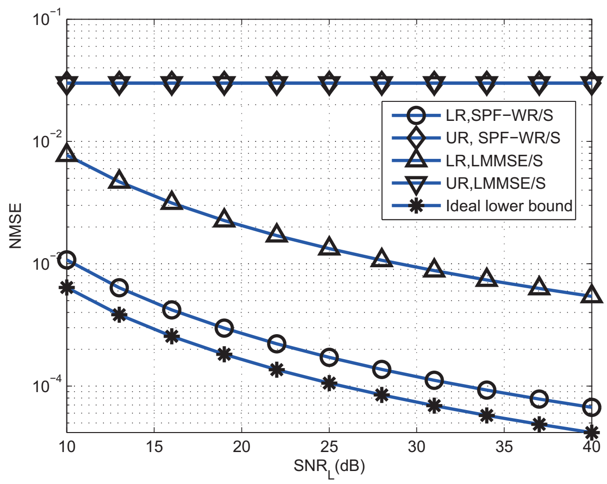
$$\begin{aligned} SNR_L &= \frac{\mathrm{E}\| \mathrm{HS}_1 \|_F^2}{\mathrm{E}\| \mathrm{E}_1 \|_F^2} = \frac{1}{\sigma_0^2}, \\ SNR_U &= \frac{\mathrm{E}\| \mathrm{GS}_1 \|_F^2}{\mathrm{E}\| \mathrm{F}_1 \|_F^2} = \frac{1}{\sigma_0^2}. \end{aligned} \tag{46}$$

In the simulation, we assume that $SNR_L = SNR_U$, and the result of each scheme is obtained by 10,000 Monte Carlo running.

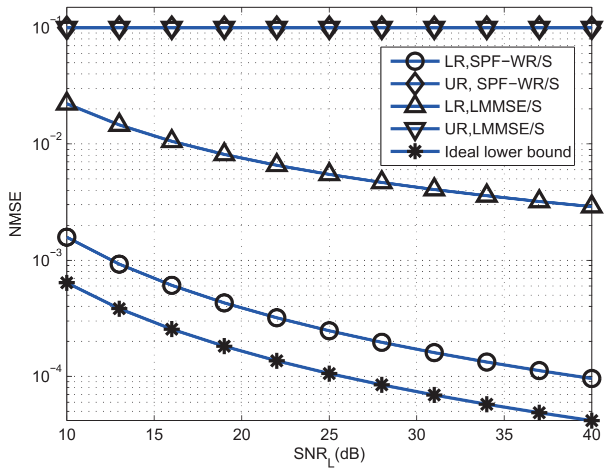
The power allocation result is depicted in Fig.2, and the NMSE versus SNR level at the LR and the UR are offered in Fig.3 (a) and (b), respectively. From Fig.3, we can see that more powers are desirable to allocate the forward training signals than the AN signals as γ increase from 0.03 to 0.1. Notice that there is a little variation to the optimal allocation between the power P_1 and σ_a^2 at all SNRs level. Thus, we can conclude that the power allocation of the proposed scheme is stable in the implement. In the DCE performance, the complex orthogonal STBC symbols [24] are offered as the forward training pilots. Fig.3 (a) and Fig.4 (b) show the results with the two schemes under the preset limit $\gamma = 0.03$ and $\gamma = 0.1$, respectively. It is found that the NMSE at the LR with the proposed DCE scheme achieves relatively higher resolution than the LMMSE/S (around 10 times higher accuracy). Moreover, we also can see that a large NMSE gap of the attainable NMSE exists between the scheme of LMMSE/S and the ideal lower bound. In conclusion, the proposed DCE scheme has a more robust performance than the LMMSE/S scheme in the noisy case.

V. Conclusion

In this paper, an improved two-way training for DCE has been proposed in the wireless MIMO systems. In particular, an alternative channel estimator, namely, WR-based semiblind channel method has been employed for enhancing the DCE performance in the noisy circumstance. Furthermore, a closed-form of the NMSE criterion has been derived to demonstrate the effectiveness of proposed scheme. Simulation results confirmed that the proposed method has a significant improvement than the LMMSE-based two-way training.



(a) $\gamma = 0.03$



(b) $\gamma = 0.1$

Fig. 3. Simulation results of NMSE performance of the LMMSE-based DCE scheme and the SPF-WR-based DCE scheme, for $N_T = 4, N_L = N_U = 2$ and $SNR_L = SNR_U$.

Acknowledgement

Rong Yu is the corresponding author of this paper. This research is partially supported by programs of NSFC (grant nos. 61370159, U1035001, U1201253, 61203117). The work of Xiangyun Zhou was supported by the Australian Research Council's Discovery Projects funding scheme (Project No. DP110102548).

Appendix A
Derivation of optimum power allocation

First, we need to deduce the range of parameter γ. To the extreme case when $y = 0$, we have

$$x \leq \frac{P_{ave}T_1}{N_T} \tag{47}$$

by using the power constraint of Eq.(43). In this case, γ has the lower bound of

$$\gamma \geq \frac{N_T\sigma_0^2}{P_{ave}T_1}. \tag{48}$$

To obtain the upper bound of γ, we consider the extreme case when $x = 0$, we have the upper bound as

$$\gamma \leq (N_T - N_L)P_{max}. \tag{49}$$

To the variable of y, it satisfies the following inequations as

$$\frac{x\gamma - \sigma_0^2}{\sigma_G^2} \leq y \leq P_{ave} - \frac{xN_T}{T_1}. \tag{50}$$

by using the constraints of optimal model (43). Besides, the inequations (50) holds if and only if

$$P_{ave} - \frac{xN_T}{T_1} \geq \frac{x\gamma - \sigma_0^2}{\sigma_G^2} \tag{51}$$

Thus, variable x satisfies that

$$\frac{\sigma_0^2}{\gamma} \leq x \leq \frac{(\sigma_G^2 P_{ave} + \sigma_0^2)T_1}{\gamma T_1 + N_T\sigma_G^2}. \tag{52}$$

From Eq.(52), the objective function is monotonically decreasing with respect to the variable y, so the optimum value can be achieved when

$$y(x) = \frac{x\gamma - \sigma_0^2}{\sigma_G^2}. \tag{53}$$

In conclusion, the optimal problem of (43) is reformulated as solving the following one variable optimum problem.

$$\begin{aligned} &\min_x \ \frac{\sigma_0^2}{x} + \frac{y(x)\sigma_0^2}{x}(1 + \frac{N_U\sigma_G^2}{N_T T_0}) \\ &s.t. \frac{\sigma_0^2}{\gamma} \leq x \leq \frac{(\sigma_G^2 P_{ave} + \sigma_0^2)T_1}{\gamma T_1 + N_T\sigma_G^2}, \end{aligned} \tag{54}$$

where

$$y(x) = \frac{x\gamma - \sigma_0^2}{\sigma_G^2}.$$

References

[1] R. M. Needham and M. D. Schrodr, “Using encryption for authentication in large networks of computers,” *Commun. ACM.* vol.21, no 12, pp. 993-999, Dec. 1978.

[2] T. Y. Wang, L. Y. Song, Z. Han, and B.L. Jiao, “Dynamic popular content distribution in vehicular networks using coalition formation games,” *IEEE Journal on Selected Areas in Commun.*, vol. 30, no. 9, pp. 538-546, Sep. 2013.

[3] R. Yu, Y. Zhang, S. Gjessing, W. Xia and K. Yang, “Toward cloud-based vehicular networks with efficient resource management,” *IEEE Network Magazine*, vol. 27, no. 5, pp. 48-55, 2013.

[4] R. Yu, Y. Zhang, Y. Liu, S. Gjessing, M. Guizani, “ Securing cognitive radio networks against primary user emulation attacks,” to appear in *IEEE Network Magazine*.

[5] R. Yu, Y. Zhang, S. Gjessing, Y. Chau, S. L. Xie, and M. Guizani, “ Cognitive radio based hierarchical communication architecture for smart grid,” *IEEE Network Magazine, special issue on Communication Infrastructures for Smart Grid* , vol. 25, no. 5, Sept./Oct. 2011.

[6] R. Yu, C. Zhang, et al., “ Hybrid spectrum access in cognitive radio based smart grid communications networks,” *IEEE System Journal, special issue on Smart Grid Communications Systems*, DOI: 10. 1109/ JSYST. 2013. 2260931, 2013.

[7] Y. Zhang, R. Yu, S. L. Xie, W. Q. Yao, Y. Xiao, and M. Guizani, “Home M2M networks: architecture, standards and QoS improvements,” *IEEE Communications Magazine*, vol. 49, no. 4, pp. 44-52, April 2011.

[8] Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. L. Xie, and S. Gjessing, “Cognitive Machine-to-Machine communications: visions and potentials for the smart grid,” *IEEE Network Magazine*, vol. 26, no. 3, pp. 6-13, May/June 2012.

[9] A. D. Wyner, “ The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355-1387, Oct. 1975.

[10] C. E. Landwhr and D. M. Goldchlag, “Secure issues in networks with internet acess,” *Proc. IEEE,* vol. 85, pp. 2034-2051, Dec. 1997.

[11] P. Gopala, L. Lai, and H. EI Gamal, “ On the secrecy capacity of fading channels,” *IEEE Trans. Inf. Theory.*, vol. 54, pp. 4687-4698, Oct. 2008.

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

[13] T. H. Chang, W. C. Chiang, Y. W. Hong, 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, 2010.

[14] C. W. Huang, T. H. Chang, X. Zhou, Y. P. Hong, “ Two-way training for discriminatory channel estimation in wireless MIMO systems,” *IEEE Trans on Signal Process.*, vol. 61, no. 10, pp. 2724-2738, 2013.

[15] X. Li and J. H. Wu, “Using antenna array redundancy and channel diversity for secure wireless transmissions,” *Journal of Communication.*, vol. 2, no. 3, May. 2007.

[16] S. Goel, R. Negi, “ Guaranteeing secrecy using artificial noise,” *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2180-2189, 2008.

[17] S. M. Kay, “Fundamentals of statistical signal processing: estimation theory,” *New Jersey: Prentice Hall International*, 1993.

[18] S. Shahbazpanahi, “Semiblind multiuser MIMO channel estimation using capon and MUSIC techniques,” *IEEE Trans. Signal Process.*, vol. 54, no. 9, Sep. 2006.

[19] A. K. Jagannatham, B. D. Rao, “ Whitening-rotation-based semiblind MIMO channel estimation,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, Mar. 2006.

[20] F. Wan, W. P. Zhu, M. N. Swamy, “ A signal perturbation free whitening-rotation-based semiblind approach for MIMO channel estimation,” *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3154-3166, 2009.

[21] H. Holma, and A. Toskala, *WCDMA for UMTS: radios access for third generation mobile communications*, New York: Wiley, pp.478, 2004.

[22] G. W. Stewart, *Perturbation theory for the sigular value decomposition. in SVD and Signal Processing (II): Algorithms, Analysis and Applications,* R. J. Vaccaro, Ed. New York: Elsevier, pp. 99-109, 1991.

[23] M. Chiang, C. W. Tan, D. P. Palomar, D. O’Neill and D. Julian, “Power control by geometric programming,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640-2651, Jul. 2007.

[24] T. Vahid, J. Hamld, A. R. Calderbank, “ Space-time block coding for wireless communications: performance results,” *IEEE Selected areas in communications*, vol. 17, no. 3, pp. 451-460, Mar. 1999.