Joint Training and Beamforming Design for Performance Discrimination Using Artificial Noise

Tsung-Hui Chang, Wei-Cheng Chiang, Y.-W. Peter Hong and Chong-Yung Chi

Institute of Communications Engineering & Department of Electrical Engineering
National Tsing Hua University, Hsinchu, Taiwan 30013
Emails: changth@mx.nthu.edu.tw, {ywhong, cychi}@ee.nthu.edu.tw

***Abstract*—Recently, in multi-antenna wireless systems, the use of artificial noise (AN) in training and data transmission phases has been respectively proposed to achieve performance discrimination between a legitimate receiver (LR) and an unauthorized receiver (UR). For data transmission, an AN-aided beamforming (ANBF) scheme has been proposed where the message is sent towards LR using beamforming while AN is imposed in the null space of LR's channel to disrupt UR's reception. For channel estimation, the so-called discriminatory channel estimation (DCE) scheme has been proposed where a multi-stage training scheme is employed and AN is imposed in the null space of the estimated LR's channel obtained in previous stages to degrade the channel estimation performance of UR. In this work, the optimal power allocation between DCE and ANBF (as well as AN in both phases) is derived with the goal of maximizing the receive signal-to-interference-plus-noise ratio (SINR) of LR subject to a constraint on the maximum achievable SINR of UR. The simulation results show that, with the joint power allocation of DCE and ANBF, the SINR at LR and UR can be effectively discriminated even when UR is equipped with more antennas than the transmitter. Moreover, it is observed that the proposed joint DCE and ANBF scheme would allocate more power to the channel estimation phase compared with that using conventional channel estimation (without considering URs) in the training phase.**

I. Introduction

The problem of discriminating the reception performance between different users in a wireless system arises in many applications, such as the discrimination between paid and unpaid subscribers in TV broadcast systems and the transmission of confidential messages in broadcast channels. We consider a wireless system where a multiple-antenna transmitter wants to communicate with a (single-antenna) legitimate receiver (LR) in the presence of an (multiple-antenna) unauthorized receiver (UR) which attempts to overhear the information sent from the transmitter. Recent results in physical layer secrecy [1] have shown that secret messages can be reliably communicated between the transmitter and LR, without being eavesdropped by UR, using physical layer coding techniques. The achievable secrecy rate is shown to depend on the performance discrimination between the two receivers. These results have motivated considerable research endeavors in devising advanced signal processing methods to achieve larger performance discrimination between the receivers [2]–[5].

In particular, two signal processing approaches have been taken to achieve this goal, one in the data transmission phase and one in the channel estimation phase, both with the help of artificial noise (AN). For data transmission, an AN-aided beamforming (ANBF) scheme has been considered in [2]–[4], where the message is sent towards LR using transmit beamforming while AN is imposed in the null space of LR's channel to disrupt UR's reception. For channel estimation, a multi-stage training scheme has been proposed in [5] where, in each stage, AN is imposed in the null space of the estimated LR's channel (which is obtained in the previous stage) to degrade UR's channel estimation performance. This scheme is referred to as the *discriminatory channel estimation* (DCE) scheme. With proper power allocation between AN and data and between AN and training, these two schemes have been shown to be effective in discriminating between LR and UR performances.

In this paper, we consider the joint use of DCE in the training phase and ANBF in the data transmission phase, and study the optimal power allocation between the training and the data transmission phases. The joint use of the two schemes can discriminate between the receivers' performances more effectively, especially when the number of antennas at UR is greater than that at the transmitter [6]. To study this problem, we first analyze the signal-to-interference-plus-noise ratio (SINR) at both LR and UR, assuming that the channel state information (CSI) at the receivers and transmitter are imperfect. The inaccuracy of the channel estimate may affect the SINR performance in the data transmission phase and, thus, power must be carefully allocated between the training phase and the data transmission phase. Based on the analytical SINR expressions, we examine the optimal power allocation between the two phases with the goal of maximizing the SINR of LR subject to an upper constraint on the SINR achievable by UR. Simulation results show that the joint DCE and ANBF scheme can effectively discriminate between the receivers' reception performances. This scheme is also shown to outperform systems that employ conventional training schemes, e.g., [7], in the channel estimation phase.

II. System Model

Consider a wireless system where a transmitter sends an information message to a legitimate receiver (LR) in the presence of an unauthorized receiver (UR) that passively overhears the message sent by the transmitter. We assume that the transmitter has $N_t > 1$ antennas, the LR has only

978-1-61284-231-8/11/$26.00

a single antenna and the UR has $N_U > 1$ antennas. Denote by $\mathbf{h} \in \mathbb{C}^{N_t}$ and $\mathbf{G} \in \mathbb{C}^{N_t \times N_U}$ the channel vector of the LR and the channel matrix of the UR, respectively. The channel is assumed to remain constant over a coherence interval that consists of a training phase with length $T_{\rm CE}$ and a data transmission phase with length $T_{\rm D}$. Specifically, the transceiver processing we consider is described as follows:

Training phase: The transmitter first sends a training sequence (which is known to both LR and UR) to enable channel estimation at both receivers. Let $\hat{\mathbf{h}}$ and $\hat{\mathbf{G}}$ denote the associated channel estimates. The LR will feedback $\hat{\mathbf{h}}$ to the transmitter for the beamforming design in the ensuing data transmission phase.

Data transmission phase: With the channel estimate $\hat{\mathbf{h}}$ fed back from LR, the transmitter employs the ANBF scheme [2] for data transmission. Here, we assume that UR does not feedback its channel and is able to intercept the channel estimate $\hat{\mathbf{h}}$ sent by LR. Both LR and UR will use their channel estimates $\hat{\mathbf{h}}$ and $\hat{\mathbf{G}}$ for data detection, respectively.

The channel estimation performances at the receivers have significant impact on their data detection performances and, thus, the training sequence in the channel estimation phase and the ANBF signal in the data transmission phase can be jointly designed to further enhance the difference between LR and UR's detection performances. Next, let us first review the ANBF and the DCE schemes mentioned above.

A. Review of AN-Aided Transmit Beamforming

In the ANBF scheme, the message is sent towards the estimated LR's direction using beamforming, and AN is imposed in the null space of the estimated LR's channel in order to interrupt UR's reception. Specifically, given the channel estimate $\hat{\mathbf{h}}$, the signal transmitted under ANBF is given by [2]

$$\boldsymbol{x}(t) = \sqrt{\sigma_s^2}\left(\frac{\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}\right)s(t) + \sqrt{\frac{\sigma_a^2}{N_t-1}}\,\mathbf{N}_{\hat{h}} \cdot \mathbf{a}(t), \quad (1)$$

for $t = 1, \ldots, T_{\rm D}$, where $s(t)$ is the information signal, with $\mathrm{E}\{|s(t)|^2\} = 1$, $\mathbf{N}_{\hat{h}} \in \mathbb{C}^{N_t \times (N_t-1)}$ is a matrix which spans the left null space of $\hat{\mathbf{h}}$ and satisfies both $\mathbf{N}_{\hat{h}}^H \mathbf{N}_{\hat{h}} = \mathbf{I}_{N_t-1}$ (the $(N_t-1) \times (N_t-1)$ identity matrix) and $\hat{\mathbf{h}}^H \mathbf{N}_{\hat{h}} = \mathbf{0}$, $\mathbf{a}(t) \in \mathbb{C}^{N_t-1}$ is an AN vector where the entries are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, and $\sigma_s^2 \geq 0$ and $\sigma_a^2 \geq 0$ represent information signal and AN powers, respectively.

The received signals of LR and UR are given by

$$y(t) = \mathbf{h}^H \boldsymbol{x}(t) + w(t), \quad (2)$$
$$\mathbf{z}(t) = \mathbf{G}^H \boldsymbol{x}(t) + \mathbf{v}(t), \quad (3)$$

respectively, where $w(t)$ is the additive noise at LR with zero mean and variance σ_w^2, and $\mathbf{v}(t)$ is the additive noise vector at UR with zero mean and covariance matrix $\sigma_v^2 \mathbf{I}_{N_U}$.

When the channel estimate is perfect, i.e., when $\hat{\mathbf{h}} = \mathbf{h}$, it follows from (1), (2), (3) and the fact that $\hat{\mathbf{h}}^H \mathbf{N}_{\hat{h}} = \mathbf{0}$ that the received signals can be written as

$$y(t) = \sqrt{\sigma_s^2}\|\hat{\mathbf{h}}\| s(t) + w(t), \quad (4)$$
$$\mathbf{z}(t) = \sqrt{\sigma_s^2}\frac{\mathbf{G}^H \hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|} s(t) + \sqrt{\frac{\sigma_a^2}{N_t-1}}\,\mathbf{G}^H \mathbf{N}_{\hat{h}} \mathbf{a}(t) + \mathbf{v}(t). \quad (5)$$

This shows that only UR will be affected by AN when the transmitter has perfect knowledge of the channel $\mathbf{h}$. However, if $\hat{\mathbf{h}} \neq \mathbf{h}$, we have $\mathbf{h}^H \mathbf{N}_{\hat{h}} \neq \mathbf{0}$ and, thus, AN will leak into LR's channel reducing the achievable SINR [8].

It is worthwhile to note that ANBF is not always effective in degrading UR's performance, especially when UR has a large number of antennas and perfectly knows its own channel $\mathbf{G}$ [6]. To see this, let us assume that $N_U \geq N_t$ and $\hat{\mathbf{G}} = \mathbf{G}$, and let us express the received signal at UR as

$$\mathbf{z}(t) = \mathbf{G}^H \mathbf{Q}_{\hat{h}} \begin{bmatrix} s(t) \\ \mathbf{a}(t) \end{bmatrix} + \mathbf{v}(t), \quad (6)$$

where

$$\mathbf{Q}_{\hat{h}} \triangleq \left[\sqrt{\sigma_s^2}\left(\frac{\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}\right) \quad \sqrt{\tfrac{\sigma_a^2}{N_t-1}}\mathbf{N}_{\hat{h}}\right]. \quad (7)$$

In this case, UR can estimate the information signal $s(t)$ simply by inverting the effective channel $\mathbf{G}^H \mathbf{Q}_{\hat{h}}$ and by taking the term corresponding to $s(t)$, e.g., by computing

$$[1\ \mathbf{0}^T]\ \mathbf{Q}_{\hat{h}}^{-1}(\mathbf{G}^H)^\dagger \mathbf{z}(t) = s(t) + [1\ \mathbf{0}^T]\ \mathbf{Q}_{\hat{h}}^{-1}(\mathbf{G}^H)^\dagger \mathbf{v}(t), \quad (8)$$

where $\mathbf{Q}_{\hat{h}}^{-1}$ and $(\mathbf{G}^H)^\dagger$ denote the inverse of $\mathbf{Q}_{\hat{h}}$ and the pseudo inverse of $\mathbf{G}^H$, respectively. Equation (8) implies that UR can be free from the interfering AN. However, if UR has an imperfect channel estimate, i.e., $\hat{\mathbf{G}} \neq \mathbf{G}$, there always exists residual AN in (8) which will still interfere with UR.

The above discussions reveal the need of a training scheme that can enable LR to estimate $\mathbf{h}$ accurately while preventing UR from obtaining an accurate estimate of $\mathbf{G}$, as described in the next subsection.

B. Review of Discriminatory Channel Estimation

The key idea of DCE is to utilize the channel estimate fed back from LR to judiciously allocate AN in the training signal to disrupt the channel estimation at UR [5].

Specifically, the transmitter will send a sequence of $K+1$ training signals $\mathbf{X}_k \in \mathbb{C}^{N_t \times T_k}$ for $k = 0, \ldots, K$, where T_k is the training length at stage k and $\sum_{k=0}^{K} T_k = T_{\rm CE}$. Let $\hat{\mathbf{h}}_{k-1}$ be the channel estimate of the LR by using $\mathbf{X}_0, \ldots, \mathbf{X}_{k-1}$. The training signals $\mathbf{X}_k$ for $k = 0, 1, \ldots, K$, are given by

$$\mathbf{X}_k = \sqrt{\frac{P_k T_k}{N_t}}\mathbf{C}_k + \left(\sqrt{\frac{\sigma_{a,k}^2}{N_t-1}}\mathbf{N}_{\hat{h}_{k-1}} \cdot \mathbf{A}_k\right)\delta[k-1], \quad (9)$$

where $\delta[\cdot]$ is the unit step function, $P_k \geq 0$ and $\sigma_{a,k}^2 \geq 0$ are the respective training data power and AN power at stage k, $\mathbf{C}_k \in \mathbb{C}^{N_t \times T_k}$ represents the training data matrix satisfying $\mathbf{C}_k \mathbf{C}_k^H = \mathbf{I}_{N_t}$, $\mathbf{N}_{\hat{h}_{k-1}} \in \mathbb{C}^{N_t \times (N_t-1)}$ is a matrix that satisfies $\mathbf{N}_{\hat{h}_{k-1}}^H \mathbf{N}_{\hat{h}_{k-1}} = \mathbf{I}_{N_t-1}$ and $\hat{\mathbf{h}}_{k-1}^H \mathbf{N}_{\hat{h}_{k-1}} = \mathbf{0}$, and

$\mathbf{A}_k \in \mathbb{C}^{(N_t-1)\times T_k}$ is an AN matrix with each entry being an i.i.d. complex Gaussian random variable with zero mean and unit variance. As seen from (9), the AN at stage k is placed in the subspace orthogonal to $\hat{\mathbf{h}}_{k-1}$ in order to alleviate the negative effect of AN on LR.

Suppose that both LR and UR have perfect knowledge of the training sequences $\{\mathbf{C}_k\}_{k=0}^K$ and perform optimal linear minimum mean squared error (LMMSE) channel estimation based on this information. We assume that:

(A1) The entries of $\mathbf{h}$ and $\mathbf{G}$ are i.i.d. with zero mean and variances σ_h^2 and σ_G^2, respectively.

The channel error correlation matrix with respect to the channel estimate $\hat{\mathbf{h}} \triangleq \hat{\mathbf{h}}_K$ can be shown to be [5]

$$\mathbf{R}_{\Delta\mathbf{h}} \triangleq \mathrm{E}\{(\mathbf{h}-\hat{\mathbf{h}})(\mathbf{h}-\hat{\mathbf{h}})^H\} = \mathrm{NMSE_L}\mathbf{I}_{N_t}, \tag{10}$$

where $\mathrm{NMSE_L} \triangleq \mathrm{Tr}(\mathbf{R}_{\Delta\mathbf{h}})/N_t = f^{(K)}$ is the normalized mean squared error (NMSE) of $\hat{\mathbf{h}}$ given by the following recursive form:

$$f^{(k)} = \left(\frac{1}{f^{(k-1)}} + \frac{P_k T_k/N_t}{f^{(k-1)}\sigma_{a,k}^2 + \sigma_w^2}\right)^{-1}, \tag{11}$$

for $k = 1, \ldots, K$ with

$$f^{(0)} = \left(\frac{1}{\sigma_h^2} + \frac{P_0 T_0}{N_t \sigma_w^2}\right)^{-1}.$$

For UR, the associated channel error correlation matrix can be shown to be

$$\mathbf{R}_{\Delta\mathbf{G}} \triangleq \mathrm{E}\{(\mathbf{G}-\hat{\mathbf{G}})(\mathbf{G}-\hat{\mathbf{G}})^H\} = N_U \mathrm{NMSE_U}\mathbf{I}_{N_t}, \tag{12}$$

where $\mathrm{NMSE_U} \triangleq \mathrm{Tr}(\mathbf{R}_{\Delta\mathbf{G}})/(N_t N_U)$ is given by [5]

$$\mathrm{NMSE_U} = \left(\frac{1}{\sigma_G^2} + \frac{P_0 T_0}{N_t \sigma_v^2} + \sum_{k=1}^{K} \frac{P_k T_k/N_t}{\sigma_{a,k}^2 \sigma_G^2 + \sigma_v^2}\right)^{-1}. \tag{13}$$

It has been shown in [5] that, with proper allocation of the training data powers $\{P_k\}_{k=0}^K$ and AN powers $\{\sigma_{a,k}^2\}_{k=1}^K$, it is possible to discriminate between the channel estimation performances of LR and UR with $\mathrm{NMSE_L} \ll \mathrm{NMSE_U}$.

III. Proposed Joint DCE and ANBF Scheme

While DCE and ANBF can be employed independently in the training phase and in the data transmission phase, we are interested in the joint optimization problem, aiming at achieving a better discrimination between LR and UR's performances. To this end, next let us analyze the performance of ANBF under the scenario where both the transmitter and receivers have imperfect CSI.

A. Performance Analysis of ANBF with Imperfect CSI at the Transmitter and Receivers

To analyze the impact of imperfect CSI, let us denote the true channels $\mathbf{h}$ and $\mathbf{G}$ as

$$\mathbf{h} = \hat{\mathbf{h}} + \Delta\mathbf{h}, \quad \mathbf{G} = \hat{\mathbf{G}} + \Delta\mathbf{G}, \tag{14}$$

where $\Delta\mathbf{h}$ and $\Delta\mathbf{G}$ are the respective channel errors. Under the LMMSE criterion and **(A1)**, $\Delta\mathbf{h}$ is of zero mean and is statistically uncorrelated with $\hat{\mathbf{h}}$ (this is also true for $\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$). For ease of analysis, we further assume that:

(A2) $\hat{\mathbf{h}}$ and $\Delta\mathbf{h}$ are statistically independent; $\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$ are statistically independent[1].

We first analyze the SINR performance of LR. According to (1), (2) and (14), the receive signal at LR can be expressed as

$$\begin{aligned} y(t) &= (\hat{\mathbf{h}} + \Delta\mathbf{h})^H \boldsymbol{x}(t) + w(t) \\ &= \sqrt{\sigma_s^2}\|\hat{\mathbf{h}}\|s(t) + \sqrt{\sigma_s^2}\frac{\Delta\mathbf{h}^H\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}s(t) \\ &\quad + \sqrt{\frac{\sigma_a^2}{N_t-1}}\Delta\mathbf{h}^H\mathbf{N}_{\hat{h}}\mathbf{a}(t) + w(t), \end{aligned} \tag{15}$$

where the first term is the signal of interest and the rest are the interference terms caused by channel estimation error, AN, and additive noise, respectively. Under **(A2)**, the SINR conditioned on the channel estimate $\hat{\mathbf{h}}$ can be shown to be

$$\begin{aligned} \mathrm{SINR_L}(\hat{\mathbf{h}}) &= \frac{\|\hat{\mathbf{h}}\|^2\sigma_s^2}{\sigma_s^2\mathrm{E}\left\{\left|\frac{\Delta\mathbf{h}^H\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}\right|^2\right\} + (\frac{\sigma_a^2}{N_t-1})\mathrm{E}\{|\Delta\mathbf{h}^H\mathbf{N}_{\hat{h}}\mathbf{a}(t)|^2\} + \sigma_w^2} \\ &= \frac{\|\hat{\mathbf{h}}\|^2\sigma_s^2}{(\sigma_s^2+\sigma_a^2)\mathrm{NMSE_L} + \sigma_w^2}. \end{aligned} \tag{16}$$

It then follows from (16) and the equality

$$\mathrm{E}\{\|\hat{\mathbf{h}}\|^2\} = N_t(\sigma_h^2 - \mathrm{NMSE_L}) \tag{17}$$

that the average SINR at LR is given by

$$\mathrm{SINR_L} = \mathrm{E}\{\mathrm{SINR_L}(\hat{\mathbf{h}})\} = \frac{\sigma_s^2 N_t(\sigma_h^2 - \mathrm{NMSE_L})}{(\sigma_s^2+\sigma_a^2)\mathrm{NMSE_L} + \sigma_w^2}. \tag{18}$$

To analyze the receive SINR at UR, let us express (3) as

$$\begin{aligned} \mathbf{z}(t) &= \sqrt{\sigma_s^2}\frac{\hat{\mathbf{G}}^H\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}s(t) + \sqrt{\sigma_s^2}\frac{\Delta\mathbf{G}^H\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}s(t) \\ &\quad + \sqrt{\frac{\sigma_a^2}{N_t-1}}\mathbf{G}^H\mathbf{N}_{\hat{h}}\mathbf{a}(t) + \mathbf{v}(t), \end{aligned} \tag{19}$$

where the first term corresponds to the information signal and the remaining terms are due to channel error, AN and additive noise, respectively. Since UR is equipped with multiple antennas, it can employ the optimal maximum SINR combining (MSINRC) to suppress the interference and noise terms in (19). Given $\hat{\mathbf{h}}$ and $\hat{\mathbf{G}}$, the output SINR of MSINRC for (19) can be computed as [9]

$$\mathrm{SINR_U}(\hat{\mathbf{h}}, \hat{\mathbf{G}}) = \sigma_s^2\left(\frac{\hat{\mathbf{h}}^H\hat{\mathbf{G}}}{\|\hat{\mathbf{h}}\|}\right)\hat{\mathbf{R}}^{-1}\left(\frac{\hat{\mathbf{G}}^H\hat{\mathbf{h}}}{\|\hat{\mathbf{h}}\|}\right), \tag{20}$$

[1]Note that **(A2)** is true when conventional channel estimation [7] is employed in the training phase and when $\mathbf{h}$ and $\mathbf{G}$ are complex Gaussian distributed since, in this case, $\hat{\mathbf{h}}$ and $\Delta\mathbf{h}$ ($\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$) are complex Gaussian distributed. However, **(A2)** does not hold true for DCE due to the presence of AN. The analytical results obtained in this subsection can be treated as an approximation for DCE nevertheless.

where

$$\hat{\mathbf{R}} = \sigma_s^2 \mathrm{E}\left\{ \frac{\Delta\mathbf{G}^H \hat{\mathbf{h}}\hat{\mathbf{h}}^H \Delta\mathbf{G}}{\| \hat{\mathbf{h}} \|^2} \;\middle|\; \hat{\mathbf{h}}, \hat{\mathbf{G}} \right\} + \left(\frac{\sigma_a^2}{N_t - 1} \right) \mathrm{E}\{\mathbf{G}^H \mathbf{N}_{\hat{\mathbf{h}}} \mathbf{N}_{\hat{\mathbf{h}}}^H \mathbf{G} \mid \hat{\mathbf{h}}, \hat{\mathbf{G}}\} + \sigma_v^2 \mathbf{I}_{N_U}. \quad (21)$$

By **(A1)** and **(A2)** and after some lengthy computations, one can further show that

$$\begin{aligned} &\mathrm{SINR}_{\mathrm{U}}(\hat{\mathbf{h}}, \hat{\mathbf{G}}) \\ &= \sigma_s^2 \left(\frac{N_t - 1}{\sigma_a^2} \right) \left(\frac{\| \hat{\mathbf{h}} \|^2}{\hat{\mathbf{h}}^H (\mathbf{I}_{N_t} + \frac{\sigma_a^2}{\beta(N_t - 1)} \hat{\mathbf{G}}\hat{\mathbf{G}}^H)^{-1} \hat{\mathbf{h}}} - 1 \right) \\ &\leq \frac{\sigma_s^2}{\beta} \mathrm{Tr}(\hat{\mathbf{G}}\hat{\mathbf{G}}^H), \end{aligned} \quad (22)$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix, $\beta = (\sigma_s^2 + \sigma_a^2)\mathrm{NMSE}_{\mathrm{U}} + \sigma_v^2$, and the inequality in (22) is due to $\mathrm{Tr}(\hat{\mathbf{G}}\hat{\mathbf{G}}^H)\mathbf{I}_{N_t} - \hat{\mathbf{G}}\hat{\mathbf{G}}^H \succeq \mathbf{0}$ (a positive semidefinite matrix). Thus an upper bound for UR's average SINR is given by

$$\begin{aligned} \mathrm{SINR}_{\mathrm{U}} = \mathrm{E}\{\mathrm{SINR}_{\mathrm{U}}(\hat{\mathbf{h}}, \hat{\mathbf{G}})\} &\leq \frac{\sigma_s^2}{\beta} \mathrm{E}\left\{ \mathrm{Tr}(\hat{\mathbf{G}}\hat{\mathbf{G}}^H) \right\} \\ &= \frac{\sigma_s^2 N_t N_U (\sigma_G^2 - \mathrm{NMSE}_{\mathrm{U}})}{(\sigma_s^2 + \sigma_a^2)\mathrm{NMSE}_{\mathrm{U}} + \sigma_v^2}, \end{aligned} \quad (23)$$

where the last equality is obtained from the fact that $\mathrm{E}\{\hat{\mathbf{G}}\hat{\mathbf{G}}^H\} = N_U(\sigma_G^2 - \mathrm{NMSE}_{\mathrm{U}})\mathbf{I}_{N_t}$.

B. Joint Power Allocation of DEC and ANBF (DCE-ANBF)

One can observe from (18) and (23) that the channel estimation performances $\mathrm{NMSE}_{\mathrm{L}}$ and $\mathrm{NMSE}_{\mathrm{U}}$ in the training phase have direct impacts on the achievable SINR performances of LR and UR, respectively. This implies that we should jointly optimize the DCE scheme in the training phase and ANBF in the data transmission phase, in order to effectively discriminate between LR and UR's SINR performances. In particular, let us consider the following joint power allocation problem:

$$\max_{\substack{\{P_k \geq 0\}_{k=0}^K, \{\sigma_{a,k}^2 \geq 0\}_{k=1}^K, \\ \sigma_s^2, \sigma_a^2 \geq 0}} \frac{\sigma_s^2 N_t (\sigma_h^2 - \mathrm{NMSE}_{\mathrm{L}})}{(\sigma_s^2 + \sigma_a^2)\mathrm{NMSE}_{\mathrm{L}} + \sigma_w^2} \quad (24a)$$

$$\begin{aligned} \text{s.t.}\ & \frac{\sigma_s^2 N_t N_U (\sigma_G^2 - \mathrm{NMSE}_{\mathrm{U}})}{(\sigma_s^2 + \sigma_a^2)\mathrm{NMSE}_{\mathrm{U}} + \sigma_v^2} \leq \gamma, \\ & P_0 T_0 + \sum_{k=1}^{K} (P_k + \sigma_{a,k}^2) T_k \\ & \qquad + (\sigma_s^2 + \sigma_a^2) T_{\mathrm{D}} \leq \mathcal{E}. \end{aligned} \quad (24b)$$

In (24), γ is a preset SINR constraint for UR, $\mathcal{E} > 0$ stands for the average energy constraint, and $\mathrm{NMSE}_{\mathrm{L}}$ and $\mathrm{NMSE}_{\mathrm{U}}$ are given in (11) and (13), respectively. As seen from (24), we aim to maximize LR's SINR [in (18)] while limiting the SINR achievable by UR [in (23)] to a value no greater than γ, under the total energy constraint (24b). The optimization problem in (24) is in general not convex and difficult to solve; however a suboptimal solution can be obtained by applying the techniques of monomial approximation and condensation method in the context of geometric programming (GP) [10]. These techniques involve solving a sequence of convex GPs and hence can be implemented efficiently. Readers are referred to [5] and [10] for the details.

To compare with the proposed joint power allocation problem in (24) which uses DCE in the training phase, we also consider the joint design problem that employs the conventional channel estimation (CCE) scheme [7] (without considering URs) in the training phase. This simple training scheme corresponds to a training signal in (9) with only $K = 0$ (i.e., no AN), and has the following NMSEs

$$\mathrm{NMSE}_{\mathrm{L}} = \left(\frac{1}{\sigma_h^2} + \frac{P_0 T_{\mathrm{CE}}}{N_t \sigma_w^2} \right)^{-1}, \ \mathrm{NMSE}_{\mathrm{U}} = \left(\frac{1}{\sigma_G^2} + \frac{P_0 T_{\mathrm{CE}}}{N_t \sigma_v^2} \right)^{-1}$$

for LR and UR, respectively. One can also consider jointly optimizing the CCE scheme and ANBF. The optimal power allocation problem for the joint CCE and ANBF (CCE-ANBF) scheme is identical to (24) except that the optimization variables reduce to $(P_0, \sigma_s^2, \sigma_a^2)$, the energy constraint in (24b) reduces to $P_0 T_{\mathrm{CE}} + (\sigma_s^2 + \sigma_a^2) T_{\mathrm{D}} \leq \mathcal{E}$, and the $\mathrm{NMSE}_{\mathrm{L}}$ and $\mathrm{NMSE}_{\mathrm{U}}$ above are used. Again, the monomial approximation and condensation method can be applied to deal with the associated joint power allocation problem.

IV. Simulation Results and Discussions

In this section, we present some simulation results to demonstrate the effectiveness of the proposed joint design formulation. The wireless system as described in Sec. II is considered with $N_t = 4$ and $N_U = 4$. The elements of the channels $\mathbf{h}$ and $\mathbf{G}$ are i.i.d. complex Gaussian distributed with zero mean and unit variance ($\sigma_h^2 = \sigma_G^2 = 1$). We assume that the coherence time is 400 which contains a training phase with length 100 ($T_{\mathrm{CE}} = 100$) and a data transmission phase with length 300 ($T_{\mathrm{D}} = 300$). The total energy constraint $\mathcal{E}$ is set to 400 ($\mathcal{E} = 400$). For DCE, we set $K = 3$, and hence the training length of each stage is 25 ($T_k = 25$ for $k = 0, \ldots, 3$). The standard solver `CVX` [11] is used to solve the GPs in the condensation method for handling (24). The information signals $s(t)$, $t = 1, \ldots, 300$, in the data transmission phase are uniformly drawn from 64-QAM signals. LR detects $s(t)$ from (15) using the maximum likelihood (ML) criterion; while UR first applies MSINRC to (19) followed by ML detection of $s(t)$. We evaluate the average symbol error rates (SERs) of the two receivers over $5,000$ channel realizations.

In the first example, we compare the performances of the DEC-ANBF scheme and the CCE-ANBF scheme in discriminating between LR and UR's reception performances. Figure 1 shows the simulation results of SERs versus $1/\sigma_v^2$ for $1/\sigma_w^2 = 24$ dB. For fair comparison, we set $\gamma = 22$ dB for DCE-ANBF, and $\gamma = 26$ dB for CCE-ANBF since, under this setting, both schemes yield similar SERs for UR (see the dashed lines in Fig. 1). One can clearly see that DCE-ANBF has a better capability in discriminating between the reception performances of LR and UR, especially when UR is in good channel conditions (i.e., $1/\sigma_v^2$ is large). For moderate to small $1/\sigma_v^2$, both schemes exhibit comparable discrimination

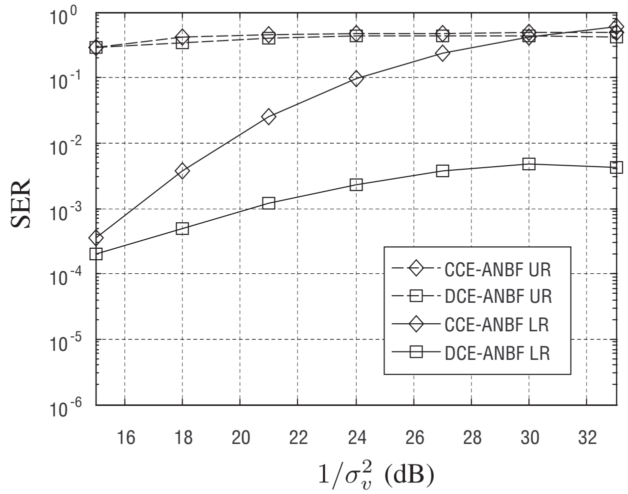

Fig. 1: SER versus $1/\sigma_v^2$ for $1/\sigma_w^2 = 24$ dB.

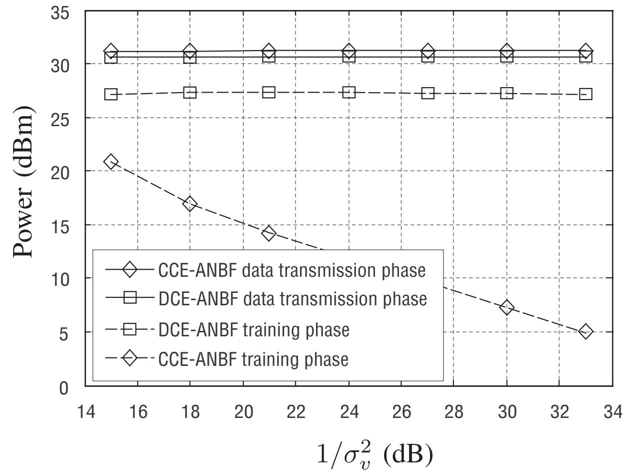

Fig. 2: Transmission powers (dBm) in the training phase and data transmission phase.

performance. In Fig. 2, we present the corresponding total training powers and the total data transmission powers of the two schemes. Firstly, we can see that DCE-ANBF allocates a slightly less data power, but a larger training power compared with CCE-ANBF. This shows that, when the employed training scheme is able to degrade UR's reception (e.g., DCE), the transmitter would like to invest more power in the training phase compared to CCE-ANBF. On the other hand, we observe from Fig. 2 that, when $1/\sigma_v^2$ increases, the training power of CCE-ANBF decreases so that more AN can be used to jam UR in the data transmission phase. By contrast, proposed DCE-ANBF allocates an almost constant power in the training phase since, in this scheme, DCE and ANBF can collaborate to cripple UR's reception using AN in both phases.

In Fig. 3, we show the simulation results of SER versus N_U for $1/\sigma_w^2 = 1/\sigma_v^2 = 24$ dB. Again, the SINR constraint γ for the UR is set to 22 dB for DCE-ANBF and set to 26 dB for CCE-ANBF. We can see from this figure that the symbol error performance of UR slightly improves with larger N_U. However, even for $N_U = 6$, the proposed DCE-ANBF scheme can still yield considerable discrimination performance.

In the future, it would be interesting to extend the joint power allocation strategy to the scenario where LR also has multiple antennas. In is anticipated that an better discrimination between LR and UR's performances can be achieved.

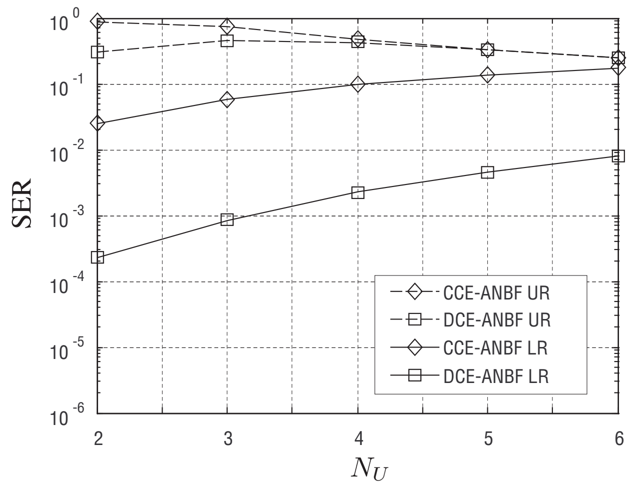

Fig. 3: SER versus N_U for $1/\sigma_w^2 = 1/\sigma_v^2 = 24$ dB.

ACKNOWLEDGMENTS

This work is supported by National Science Council (NSC), R.O.C., under grants NSC 98-2219-E-007-003, NSC 98-2219-E-007-005 and NSC-99-2221-E-007-052-MY3.

REFERENCES

[1] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE. Trans. Inform. Theory*, vol. 55, no. 7, pp. 3088–3104, July 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] A. L. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in *Proc. IEEE ICASSP*, Taipei, Taiwan, April 19-24, 2009, pp. 2437–2440.

[4] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Joint transmit beamforming and artificial noise design for QoS discrimination in wireless downlink," in *Proc. IEEE ICASSP*, Dallas, Texas, USA, March 14-19, 2010, pp. 2562–2565.

[5] 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).

[6] X. Li and J. Hwu, "Using antenna array redundancy and channel diversity for secure wireless transmissions," *Journal of Commun.*, vol. 2, no. 3, pp. 24–32, May 2007.

[7] T. F. Wong and B. Park, "Training sequence optimization in MIMO systems with colored interference," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1939–1947, Nov. 2004.

[8] Y.-L. Liang, Y.-S. Wang, T.-H. Chang, Y.-W. Hong, and C.-Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy using artificial noise," in *Proc. IEEE ISIT*, Seoul, Korea, June 28-July 3, 2009, pp. 2351–2355.

[9] D. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[10] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, pp. 67–127, April 2007.

[11] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," http://stanford.edu/∼boyd/cvx, June 2009.