OPTIMAL TRANSMISSION STRATEGY FOR OUTAGE RATE MAXIMIZATION IN MISO FADING CHANNELS WITH TRAINING

Kun-Yu Wang[*]*, Tsung-Hui Chang*[*]*, Wing-Kin Ma*[†]*, and Chong-Yung Chi*[*]

[*] Institute of Commun. Eng. & Department of Elect. Eng.
National Tsing Hua University,
Hsinchu, Taiwan 30013
E-mail: {kunyuwang7, tsunghui.chang}@gmail.com, cychi@ee.nthu.edu.tw

[†] Department of Electronic Eng.
Chinese University of Hong Kong,
Shatin, Hong Kong
E-mail: wkma@ieee.org

ABSTRACT

In this paper, we consider a single-user multiple-input single-output (MISO) fading channel with training, and investigate optimal training and data transmission strategies for outage rate maximization. The receiver obtains instantaneous channel estimates through training; while the transmitter knows only the statistical information of the channel. We present analytical, closed-form solutions for the optimal training power and optimal data transmit covariance matrix. In particular, explicit numbers of antennas required for optimal data transmission are analyzed. Numerical results are presented to validate our analysis.

Index Terms— multiple-input single-output, training design, transmitter design, outage rate maximization.

1. INTRODUCTION

Multiple transmit antennas can improve the capacity of wireless fading channels, provided that the receiver knows the channel state information (CSI). To enable the receiver to learn the CSI, the transmitter has to send training signals before data transmission. Given a total energy constraint for training and data transmission, it is of great importance to investigate optimal training and data transmission strategies for maximizing the system throughput [1].

In this paper, we consider a single-user wireless multiple-input single-output (MISO) system. The channel between the transmitter and the receiver is assumed to be independent and identically distributed (i.i.d.) Rayleigh faded, and remains unchanged during each transmission block. Each transmission block consists of a training phase, which enables channel estimation at the receiver, and a data transmission phase. The transmitter is assumed to have no instantaneous CSI. Under the same system setup but with multiple antennas at the receiver, optimal training and data transmission designs for ergodic rate maximization have been studied in [1]. In contrast to the ergodic performance, outage rate performance is more suitable for delay-limited applications such as voice and video communications. The data transmission strategies for optimal outage rate performance is, however, quite different from that for optimal ergodic rate performance, and is more difficult to analyze in general. In particular, assuming perfect CSI at the receiver, it was first observed in [2] and later proved in [3] that the optimal strategy for outage rate performance is to use only a fraction of the total number of antennas for data transmission, in contrast to its ergodic counterpart where using all antennas is always optimal [2, 4]. However, analytic conditions

for the required number of antennas for optimal data transmission are still unknown in general [3].

Our focus in this paper is to study the joint training and data transmission design problem for optimizing the outage rate performance. In particular, we aim to jointly optimize the training power and the data transmit covariance matrix in order to maximize the outage rate under a total energy constraint. It turns out that the optimal training power for the outage rate maximization problem is identical to that for the ergodic rate maximization problem studied in [1] and has closed-form solutions. For the optimal data transmit covariance matrix, we also present analytic closed-form solutions with explicit number of antennas required for optimal data transmission. Our analytic results extend upon the results in [3] by applying theorems on the extremal probabilities of quadratic forms of Gaussian random variables in [5]. Numerical results are presented to verify our theoretical claims.

2. SIGNAL MODEL AND PROBLEM STATEMENT

Consider a single-user MISO wireless system, where the transmitter is equipped with N_t antennas. We assume a block fading channel. Specifically, the channel vector between the transmitter and the receiver, denoted by $\boldsymbol{h} \in \mathbb{C}^{N_t}$, is assumed to be circularly symmetric complex Gaussian distributed with zero mean and covariance matrix $\sigma_h^2 \boldsymbol{I}_{N_t}$ (an $N_t \times N_t$ identity matrix), i.e., $\boldsymbol{h} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_h^2 \boldsymbol{I}_{N_t})$, and the coefficients of $\boldsymbol{h}$ remain static in one transmission block but can vary from block to block. Each transmission block consists of two phases – a training phase with length T_c, followed by a data transmission phase with length T_d.

In the training phase, the transmitter sends a training signal to enable channel estimation at the receiver. Assume that the transmitter employs the optimal training scheme in [6], with P_c being the average training power, and that the receiver performs linear minimum mean squared error (LMMSE) channel estimation. Denote $\hat{\boldsymbol{h}} \in \mathbb{C}^{N_t}$ as the LMMSE channel estimate, and $\boldsymbol{e}_c = \boldsymbol{h} - \hat{\boldsymbol{h}}$ as the estimation error vector. Both $\boldsymbol{e}_c$ and $\hat{\boldsymbol{h}}$ are complex Gaussian distributed with zero mean and covariance matrices $\sigma_e^2 \boldsymbol{I}_{N_t}$ and $(\sigma_h^2 - \sigma_e^2)\boldsymbol{I}_{N_t}$, respectively [6], where

$$\sigma_e^2 = \left(\frac{1}{\sigma_h^2} + \frac{T_c P_c}{N_t \sigma^2} \right)^{-1}, \qquad (1)$$

and σ^2 is the additive Gaussian noise power at the receiver.

The receiver will use the channel estimate $\hat{\boldsymbol{h}}$ for data reception in the data transmission phase. Since it is difficult to obtain exact formulas for the channel capacity in the presence of channel estimation error [1, 7], we consider achievable lower bounds that possess

closed-form expressions. In particular, assuming that the input data signal $\boldsymbol{x}(t) \in \mathbb{C}^{N_t}$ has zero mean and covariance matrix $\boldsymbol{Q} \succeq \boldsymbol{0}$ (positive semidefinite), an achievable rate is given by [1, 7]

$$\bar{T}_d \log_2 \left(1 + \frac{\hat{\boldsymbol{h}}^H \boldsymbol{Q} \hat{\boldsymbol{h}}}{\sigma_e^2 \mathrm{Tr}(\boldsymbol{Q}) + \sigma^2} \right) \text{ bits/sec/Hz,} \qquad (2)$$

where $\bar{T}_d = T_d/(T_c + T_d)$, and $\mathrm{Tr}(\boldsymbol{Q})$ is the trace of matrix $\boldsymbol{Q}$.

The scenario under consideration is that there is no instantaneous channel estimate feedback from the receiver, and that the transmitter knows only the statistical information of $\hat{\boldsymbol{h}}$. Under such circumstances, our goal is to jointly optimize the training power P_c and the transmit covariance matrix $\boldsymbol{Q}$ such that the ρ-outage rate, i.e., the achievable rate for which the probability of rate outage is no larger than $\rho \in [0, 1)$, can be maximized. Mathematically, this outage constrained design can be formulated as the following optimization problem:

$$\max_{\boldsymbol{Q} \in \mathbb{H}^{N_t}, P_c, R \in \mathbb{R}} R \qquad (3a)$$

$$\text{s.t. } \mathrm{Prob} \left\{ \bar{T}_d \log_2 \left(1 + \frac{\hat{\boldsymbol{h}}^H \boldsymbol{Q} \hat{\boldsymbol{h}}}{\sigma_e^2 \mathrm{Tr}(\boldsymbol{Q}) + \sigma^2} \right) < R \right\} \leq \rho, \qquad (3b)$$

$$T_c P_c + \mathrm{Tr}(\boldsymbol{Q}) T_d \leq \mathcal{E}_{max}, \qquad (3c)$$

$$P_c \geq 0, \ R \geq 0, \ \boldsymbol{Q} \succeq \boldsymbol{0}, \qquad (3d)$$

where $\mathcal{E}_{max} > 0$ is the maximum energy constraint. The joint training and data transmission design problem in (3) is difficult to handle due to the probability constraint (3b). In the next section, we show how explicit solutions of P_c, $\boldsymbol{Q}$ and R for problem (3) can be analytically obtained.

3. OPTIMAL TRANSMISSION STRATEGY

As will be seen soon, the optimal P_c and $\boldsymbol{Q}$ can be obtained separately; they are respectively presented in the subsequent two subsections.

3.1. Optimal Training Power Design

Let us express

$$\boldsymbol{Q} = \mathrm{Tr}(\boldsymbol{Q})\tilde{\boldsymbol{Q}} \quad \text{and} \quad \hat{\boldsymbol{h}} = \left(\sigma_h^2 - \sigma_e^2 \right)^{1/2} \hat{\boldsymbol{u}} \qquad (4)$$

where $\tilde{\boldsymbol{Q}} \succeq \boldsymbol{0}$, $\mathrm{Tr}(\tilde{\boldsymbol{Q}}) = 1$, and $\hat{\boldsymbol{u}} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_t})$ are power normalized counterparts of $\boldsymbol{Q}$ and $\hat{\boldsymbol{h}}$, respectively. Note that the energy constraint (3c) will hold with equality when the optimal R is achieved; otherwise one can always obtain a higher rate R by scaling up $\mathrm{Tr}(\boldsymbol{Q})$. Hence one can write (3c) as

$$\mathrm{Tr}(\boldsymbol{Q}) = (\mathcal{E}_{max} - T_c P_c)/T_d. \qquad (5)$$

By (4) and (5), one can express problem (3) as

$$\max_{\substack{P_c, R \geq 0, \\ \tilde{\boldsymbol{Q}} \succeq \boldsymbol{0}, \mathrm{Tr}(\tilde{\boldsymbol{Q}}) = 1}} R \qquad (6a)$$

$$\text{s.t. } \mathrm{Prob} \left\{ \bar{T}_d \log_2(1 + \alpha(P_c)\hat{\boldsymbol{u}}^H \tilde{\boldsymbol{Q}} \hat{\boldsymbol{u}}) < R \right\} \leq \rho, \quad (6b)$$

$$\alpha(P_c) = \frac{(\sigma_h^2 - \sigma_e^2)(\mathcal{E}_{max} - T_c P_c)/T_d}{\sigma_e^2 (\mathcal{E}_{max} - T_c P_c)/T_d + \sigma^2}. \qquad (6c)$$
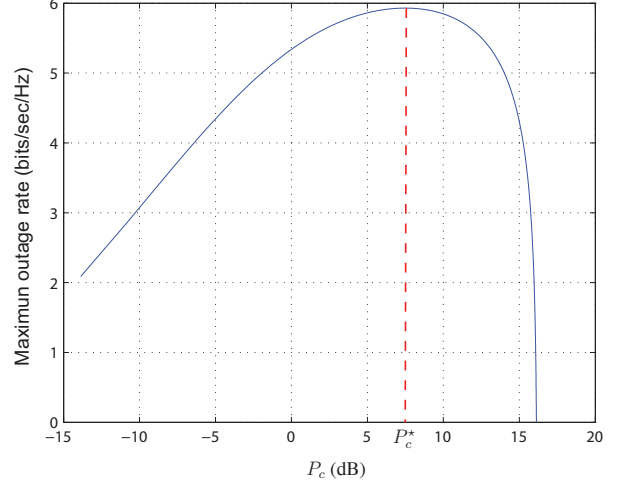

Fig. 1: Maximum outage rate versus training power. The optimal training power $P_c^\star = 7.499$ dB.

It is not difficult to verify from (6) that the optimal P_c, denoted by $P_c^\star$, is the one that maximizes $\alpha(P_c)$. Interestingly, the optimal training power for the rate outage maximization problem (3) is identical to that in [1] for ergodic rate maximization. A closed-form solution of $P_c^\star$ has been derived in [1, Theorem 2] and is summarized in the following lemma:

Lemma 1 *The optimal training power $P_c^\star$ to problem* (3) *is given by*

$$P_c^\star = \arg \max_{P_c \geq 0} \alpha(P_c) = \begin{cases} (1 - \varphi_1)\mathcal{E}_{max}/T_c, & T_d > N_t, \\ \mathcal{E}_{max}/(2T_c), & T_d = N_t, \\ (1 - \varphi_2)\mathcal{E}_{max}/T_c, & T_d < N_t, \end{cases} \quad (7)$$

where $\varphi_1 = \xi - \sqrt{\xi(\xi - 1)}$, $\varphi_2 = \xi + \sqrt{\xi(\xi - 1)}$, and

$$\xi = \frac{\sigma^2 N_t + \sigma_h^2 \mathcal{E}_{max}}{\sigma_h^2 \mathcal{E}_{max} \left(1 - \frac{N_t}{T_d} \right)}. \qquad (8)$$

While the rate outage maximization problem in (3) and the ergodic rate maximization problem in [1] have the same strategy in allocating the training and data powers, as one will see later, the outage rate maximization problem (3) can have a very different data transmission strategy from its ergodic counterpart in [1].

Before proceeding to the optimal $\tilde{\boldsymbol{Q}}$, let us present a simulation example demonstrating the importance of optimal training power design. Figure 1 displays the maximum outage rate versus training power P_c, for $N_t = 4$, $T_c = N_t$, and $T_d = 40N_t$. The channel variance $\sigma_h^2 = 1$, noise variance $\sigma^2 = 0.01$, energy constraint $\mathcal{E}_{max} = 164$ (i.e., the average transmit power $\mathcal{E}_{max}/(T_c + T_d) = 1$), and outage probability $\rho = 0.1$. The normalized data covariance matrix $\tilde{\boldsymbol{Q}} = (1/N_t)\boldsymbol{I}_{N_t}$, which will be shown to be the optimal strategy under this simulation setting. One can observe from Fig. 1 that an improper training power can result in dramatic rate reduction.

3.2. Optimal Transmit Covariance Matrix Design

Given the optimal training power $P_c^\star$, (6) can be simplified to

$$\max_{\substack{R \geq 0, \\ \tilde{\boldsymbol{Q}} \succeq \boldsymbol{0}, \mathrm{Tr}(\tilde{\boldsymbol{Q}}) = 1}} R \qquad (9a)$$

$$\text{s.t. } \mathrm{Prob} \left\{ \bar{T}_d \log_2(1 + \alpha(P_c^\star)\hat{\boldsymbol{u}}^H \tilde{\boldsymbol{Q}} \hat{\boldsymbol{u}}) < R \right\} \leq \rho. \quad (9b)$$

We can observe from problem (9) that the inequality constraint in (9b) must hold with equality when the optimal $\tilde{Q}$ and R are achieved. As a result, the optimal $\tilde{Q}$ must be the one that minimizes the probability function in (9b) since it always admits a higher outage rate R. It follows from the two observations that

$$P_{min}(\gamma^\star) \triangleq \min_{\text{Tr}(\tilde{Q})=1, \tilde{Q} \succeq \mathbf{0}} \text{Prob}\left\{\hat{u}^H \tilde{Q} \hat{u} \leq \gamma^\star\right\} = \rho, \quad (10)$$

where

$$\gamma^\star \triangleq \frac{2^{R^\star/\bar{T}_d} - 1}{\alpha(P_c^\star)}, \quad (11)$$

and $R^\star$ denotes the optimal outage rate. According to (10) and (11), once we can fully characterize $P_{min}(\cdot)$, then the optimal outage rate $R^\star$ can be simply obtained as

$$R^\star = \bar{T}_d \log_2(1 + \alpha(P_c^\star) P_{min}^{-1}(\rho)), \quad (12)$$

where $P_{min}^{-1}(\cdot)$ is the inverse function of $P_{min}(\cdot)$.

We therefore focus on analyzing the following function

$$P_{min}(x) = \min_{\text{Tr}(\tilde{Q})=1, \tilde{Q} \succeq \mathbf{0}} \text{Prob}\left\{\hat{u}^H \tilde{Q} \hat{u} \leq x\right\}, \quad (13)$$

where $\hat{u} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_{N_t})$. To this end, consider the eigenvalue decomposition of $\tilde{Q} = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U}^H$, where $\boldsymbol{U} \in \mathbb{C}^{N_t \times N_t}$ is a unitary matrix and $\tilde{\boldsymbol{\Lambda}} \in \mathbb{R}^{N_t \times N_t}$ is a diagonal matrix with eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{N_t} \geq 0$ being the diagonal elements. Since Gaussian random variables are invariant with unitary transformation, the probability function in (13) can be written as $\text{Prob}\{\hat{u}^H\tilde{\boldsymbol{\Lambda}}\hat{u} \leq x\} = \text{Prob}\{\sum_{i=1}^{N_t} \tilde{\lambda}_i|\hat{u}_i|^2 \leq x\}$, where $\hat{u}_i$ denotes the ith entry of $\hat{u}$. Further define $v_i \sim \mathcal{N}(0,1)$, $i = 1, \ldots, 2N_t$, as independent real-valued Gaussian random variables and let

$$\lambda_{2k} = \lambda_{2k-1} = \tilde{\lambda}_k/2, \ k = 1, \ldots, N_t. \quad (14)$$

By the fact of $|\hat{u}_k|^2 = (v_{2k}^2 + v_{2k-1}^2)/2$, $k = 1, \ldots, N_t$, $P_{min}(x)$ in (13) can be expressed as

$$P_{min}(x) = \min_{\substack{\sum_{i=1}^{2N_t} \lambda_i = 1, \\ \lambda_{2k}=\lambda_{2k-1} \geq 0, k=1,\ldots,N_t}} \text{Prob}\left\{\sum_{i=1}^{2N_t} \lambda_i v_i^2 \leq x\right\}. \quad (15)$$

It has been shown in [3] that $P_{min}(x)$ is the cumulative distribution function (CDF) of a central chi-square random variable with certain even number degrees of freedom (DoFs), say $1 \leq d^\star \leq 2N_t$; i.e.,

$$P_{min}(x) = \text{Prob}\left\{\frac{1}{d^\star}\sum_{i=1}^{d^\star} v_i^2 \leq x\right\}. \quad (16)$$

However, it is still not clear how to analytically determine $d^\star$ in general. Here we resolve this issue by presenting an exact expression of $P_{min}(x)$. Our analysis relies on the fact of (16) and the theorems on the extremal probabilities of quadratic forms of Gaussian random variables in [5]. To elaborate upon this, we first define some useful notations:

Definition 1 *For positive integers $d \geq 1$ and $\ell \geq 1$, $x(d, d+\ell)$ represents the point at which the CDFs $F_d(x) \triangleq \text{Prob}\left\{d^{-1}\chi_d^2 \leq x\right\}$ and $F_{d+\ell}(x)$ intersect, where χ_d^2 denotes the central chi-square random variable with d DoFs. The notation*

$$p(d, d+\ell) = F_d(x(d, d+\ell)) = F_{d+\ell}(x(d, d+\ell)) \quad (17)$$

represents the corresponding probability value of $x(d, d+\ell)$.

Moreover, $x(d, d+\ell)$ and $p(d, d+\ell)$ have the following monotonic properties:

Lemma 2 [5] *i) $x(d, d+\ell)$ is unique, larger than one, and decreases to one as d increases; ii) $p(d, d+\ell)$ is greater than 0.5, and decreases to 0.5 as d increases.*

By Lemma 2, we can define $x(0, \ell) \triangleq \infty$ and $p(0, \ell) \triangleq 1$ for any ℓ. Precise values of $x(d, d+\ell)$ and $p(d, d+\ell)$ for $d \geq 1$ and $\ell \geq 1$ can be computed numerically. Using the notations in Definition 1 and Lemma 2, we prove in Section 4 the following proposition on the solution of (15).

Proposition 1 *$P_{min}(x)$ in (15) is continuous and monotonically increasing in x, and can be explicitly expressed as*

$$P_{min}(x) = \begin{cases} F_{2n}(x), & \forall x \in [x(2n, 2n+2), x(2n-2, 2n)), \\ & n = 1, \ldots, N_t - 1, \\ F_{2N_t}(x), & \forall x \in [0, x(2N_t - 2, 2N_t)). \end{cases}$$

Proposition 1 shows that $P_{min}(x)$ is composed of $F_{2n}(x)$, $n = 1, \ldots, N_t$ in a piece-wise manner. Combining Proposition 1, Definition 1 and (10), and by the monotonicity of $P_{min}(x)$, we can obtain explicit expression of $P_{min}^{-1}(\rho)$ as

$$P_{min}^{-1}(\rho) = \begin{cases} F_{2n}^{-1}(\rho), & \forall \rho \in [p(2n, 2n+2), p(2n-2, 2n)), \\ & n = 1, \ldots, N_t - 1, \\ F_{2N_t}^{-1}(\rho), & \forall \rho \in [0, p(2N_t - 2, 2N_t)). \end{cases} \quad (18)$$

Then the optimal outage rate $R^\star$ of problem (3) can be obtained by (18) and (12). Equation (18) also implies that the optimal DoFs in (16) is given by

$$d^\star = \begin{cases} 2n, & \forall \rho \in [p(2n, 2n+2), p(2n-2, 2n)), \\ & n = 1, \ldots, N_t - 1, \\ 2N_t, & \forall \rho \in [0, p(2N_t - 2, 2N_t)). \end{cases} \quad (19)$$

With $d^\star$, the optimal $\tilde{Q}$ can be obtained as

$$\tilde{Q}^\star = \boldsymbol{U}^\star \boldsymbol{\Lambda}(d^\star/2)(\boldsymbol{U}^\star)^H, \quad (20)$$

where $\boldsymbol{U}^\star \in \mathbb{C}^{N_t \times N_t}$ can be an arbitrary unitary matrix and $\boldsymbol{\Lambda}(d^\star/2) \in \mathbb{R}^{N_t \times N_t}$ is a diagonal matrix with the first $d^\star/2$ diagonal elements being nonzero and equal to $2/d^\star$ (due to $\text{Tr}(\tilde{Q}) = 1$).

It is interesting to note from (19) and (20) that it is not necessary to use all the DoFs for optimal data transmission, especially when ρ is high; this result is in strong contrast to that in [2,4] for ergodic rate maximization where it is shown that using all the available DoFs, i.e., $\tilde{Q}^\star = (1/N_t)\boldsymbol{I}_{N_t}$, is always optimal. One can also see from (19) and (20) that for $\rho \geq p(2, 4) = 0.7153$, $d^\star = 2$ and thus using only one antenna for data transmission is sufficient to be optimal. Conversely, when $\rho < p(2N_t - 2, 2N_t)$ (where $p(2N_t - 2, 2N_t) > 0.5$), $d^\star = 2N_t$ and hence the optimal transmission strategy is to equally allocate powers to all N_t antennas. For the other cases, it is sufficient to use only a fraction of N_t antennas for optimal data transmission.

In Fig. 2, we present the simulation results of the outage rate achieved by the optimal strategy in (19) and (20). The maximum achievable outage rates of the single-antenna transmission strategy (which corresponds to $\tilde{Q} = \boldsymbol{\Lambda}(1)$) and the all-antenna transmission strategy (which corresponds to $\tilde{Q} = (1/N_t)\boldsymbol{I}_{N_t}$) are also presented. We can see from this figure that the numerical results are consistent with our analytical results presented in this subsection.
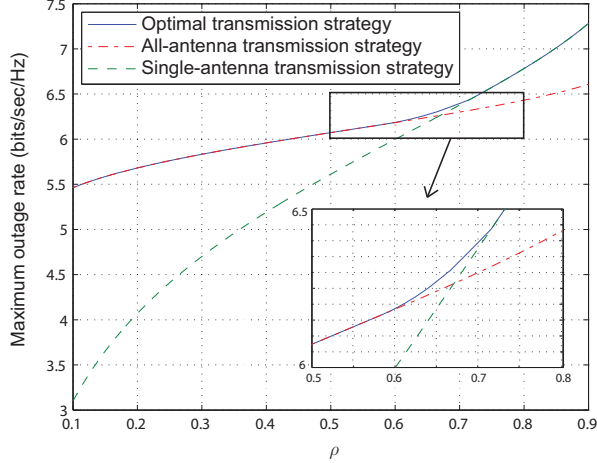
Fig. 2: Maximum outage rate versus outage probability ρ of different transmission strategies. $N_t = 10$, $T_c = N_t$, $T_d = 40N_t$, $\sigma_h^2 = 1$, $\sigma^2 = 0.01$, and $\mathcal{E}_{max}/(T_c + T_d) = 1$.

4. PROOF OF PROPOSITION 1

We first review some important analysis results in [5]. Specifically, in [5], it was shown that the following function

$$\bar{P}_{min}(x) \triangleq \min_{\substack{\lambda_i \geq 0,\ i=1,\ldots,2N_t, \\ \sum_{i=1}^{2N_t} \lambda_i = 1}} \text{Prob}\left\{\sum_{i=1}^{2N_t} \lambda_i v_i^2 \leq x\right\}, \quad (21)$$

where $v_i \sim \mathcal{N}(0,1)$, is continuous and monotonically increasing, and has a closed-form expression as

$$\bar{P}_{min}(x) = \begin{cases} F_d(x), & \forall x \in [x(d, d+1), x(d-1, d)), \\ & d = 1, \ldots, 2N_t - 1, \\ F_{2N_t}(x), & \forall x \in [0, x(2N_t - 1, 2N_t)). \end{cases} \quad (22)$$

Figure 3 illustrates $\bar{P}_{min}(x)$. However, the above result is not directly applicable to (15) since the latter has the additional constraint of $\lambda_{2k} = \lambda_{2k-1} \geq 0$, $k = 1, \ldots, N_t$. By the fact of (16), this constraint is equivalent to limiting the optimal $d^\star$ to be an even number. Therefore, for the case of $x \in [x(2n, 2n+1), x(2n-1, 2n))$, where $n \geq 1$ is an integer, we can directly obtain from (22) that $d^\star = 2n$; i.e.,

$$P_{min}(x) = F_{2n}(x) \ \forall x \in [x(2n, 2n+1), x(2n-1, 2n)). \quad (23)$$

Now let us consider the case of $x \in [x(2n-1, 2n), x(2n-2, 2n-1))$. We have to choose $1 \leq d \leq 2N_t$ among even numbers such that $F_d(x)$ is minimum in the interval $[x(2n-1, 2n), x(2n-2, 2n-1))$. We will use the following lemma:

Lemma 3 [5, Proposition 1′] *For any integers $d_2 > d_1 > 0$, there exists a unique point $x(d_1, d_2) \in [x(d_2-1, d_2), x(d_1, d_1+1)]$ such that $F_{d_2}(x) \geq F_{d_1}(x)$ for $x \geq x(d_1, d_2)$, and $F_{d_2}(x) < F_{d_1}(x)$ for $x \in (0, x(d_1, d_2))$.*

Let $\tilde{d} > 2n$ be an even number. By Lemma 3, there exists a point $x(2n, \tilde{d}) \in [x(\tilde{d}-1, \tilde{d}), x(2n, 2n+1))$ such that $F_{\tilde{d}}(x) \geq F_{2n}(x)$ for $x \geq x(2n, \tilde{d})$, and $F_{\tilde{d}}(x) < F_{2n}(x)$ for $x \in (0, x(2n, \tilde{d}))$, i.e., $x(2n, \tilde{d})$ is the crossing point. Therefore, $F_{2n}(x)$ is the minimum
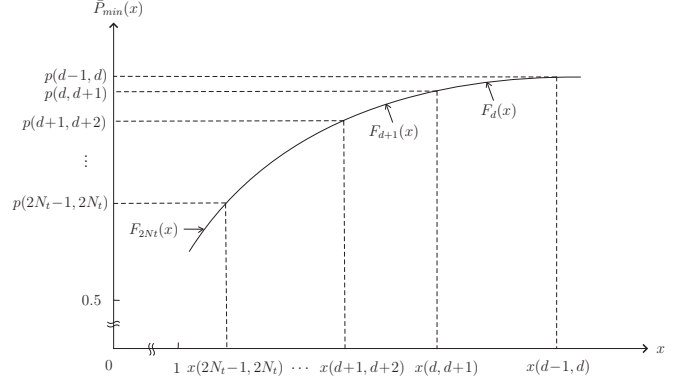

Fig. 3: Illustration of $\bar{P}_{min}(x)$ in (22).

CDF in the interval $[x(2n-1, 2n), x(2n-2, 2n-1))$ compared to any other $F_{\tilde{d}}(x)$ with even number $\tilde{d} > 2n$. Analogously, one can show that $F_{2n-2}(x)$ is the minimum CDF in the interval $[x(2n-1, 2n), x(2n-2, 2n-1))$ compared to any other $F_{\tilde{d}}(x)$ with even number $\tilde{d} < 2n - 2$.

What remains is to choose between $F_{2n}(x)$ and $F_{2n-2}(x)$. Applying Lemma 3 to $F_{2n}(x)$ and $F_{2n-2}(x)$ again, we can obtain that there exists a point $x(2n-2, 2n) \in [x(2n-1, 2n), x(2n-2, 2n-1))$ such that $F_{2n}(x) \geq F_{2n-2}(x)$ for $x \geq x(2n-2, 2n)$, and $F_{2n}(x) < F_{2n-2}(x)$ for $x \in (0, x(2n-2, 2n))$. This implies that

$$P_{min}(x) = F_{2n}(x) \ \forall x \in [x(2n-1, 2n), x(2n-2, 2n)). \quad (24)$$

By using similar argument to the interval $[x(2n+1, 2n+2), x(2n, 2n+1))$, we can also obtain

$$P_{min}(x) = F_{2n}(x) \ \forall x \in [x(2n, 2n+2), x(2n, 2n+1)). \quad (25)$$

Combining (23), (24), and (25), we conclude with

$$P_{min}(x) = F_{2n}(x) \ \forall x \in [x(2n, 2n+2), x(2n-2, 2n)), \quad (26)$$

for $n = 1, \ldots, N_t - 1$, and

$$P_{min}(x) = F_{2N_t}(x) \ \forall x \in [0, x(2N_t - 2, 2N_t)). \quad (27)$$

The proof is thus completed. ∎

5. REFERENCES

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

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Bell Labs Tech. J.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[3] E. A. Jorswieck and H. Boche, "Outage probability in multiple antenna systems," *European Trans. Telecommun.*, vol. 18, no. 3, pp. 217–233, Apr. 2007.

[4] ——, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[5] G. J. Székely and N. K. Bakirov, "Extremal probabilities for Gaussian quadratic forms," *Probability Theory and Related Fields*, vol. 126, no. 2, pp. 184–202, Jun. 2003.

[6] 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.

[7] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.