On the Impact of Quantized Channel Feedback in Guaranteeing Secrecy with Artificial Noise: The Noise Leakage Problem

Shih-Chun Lin, *Member, IEEE*, Tsung-Hui Chang, *Member, IEEE*, Ya-Lan Liang, Y.-W. Peter Hong, *Member, IEEE*, and Chong-Yung Chi, *Senior Member, IEEE*

***Abstract*—The impact of quantized channel direction information (CDI) on the achievable secrecy rate is studied for multiple antenna wiretap channels. By assuming that the eavesdropper's channel is unknown at the transmitter, we adopt the transmission scheme where artificial noise (AN) is imposed in the null space of the legitimate receiver's channel to disrupt the eavesdropper's reception. It has been shown that, in the ideal case where perfect CDI is available at the transmitter, the achievable secrecy rate can be made arbitrarily large by increasing the transmission power. However, when only quantized CDI is available, the AN that was originally intended to jam the eavesdropper may now leak into the legitimate receiver's channel, causing significant secrecy rate loss. For a given number of feedback bits B and transmission power P, we derive the optimal power allocation among the message-bearing signal and the AN to maximize the secrecy rate under AN leakage. We show that, when B is sufficiently large, one should allocate power evenly among the message-bearing signal and the AN; whereas when B is small, one should be more conservative in allocating power to the AN. Moreover, by showing that the achievable secrecy rate under quantized CDI is bounded by a constant, we derive a scaling law between B and P that is necessary to maintain a constant secrecy rate loss compared to the perfect CDI case. The scaling of B is shown to be logarithmic of P. These results are first derived for the multiple-input single-output single–antenna-eavesdropper scenario and are later extended to the multiple-input multiple-output multiple-antenna-eavesdropper case. Numerical simulations are provided to verify our theoretical claims.**

***Index Terms*—Wiretap channels, secrecy, MIMO, beamforming, quantized channel.**

I. Introduction

THE notion of physical-layer secrecy was first introduced by Wyner in [1], where the maximum achievable secrecy rate between the transmitter and a legitimate receiver is examined subject to a secrecy constraint on the information attainable by an eavesdropper. Under the perfect secrecy constraint where the eavesdropper is not allowed to infer any information from its received signal, a non-zero secrecy capacity in a static channel can only be achieved when the legitimate receiver has a better channel condition than the eavesdropper. Yet, this can be overcome in wireless environments by exploiting the time-varying characteristic of fading channels [2], [3]. Further enhancements are attainable by employing multiple antennas at the transceivers, e.g., in [4]–[6]. However, most of these works rely on perfect knowledge of the legitimate receiver's and the eavesdropper's channels. To guarantee secrecy without knowing the eavesdropper channel, the work in [5] proposed the use of artificial noise (AN) in the null space of the legitimate channel to disrupt the eavesdropper's reception. The secret message is then beamformed towards the legitimate receiver on top of the AN. With perfect knowledge of the legitimate receiver's channel direction information (CDI) at the transmitter, it has been shown that the secrecy rate achievable by using AN can be made arbitrarily large by increasing the transmission power. However, this may not be the case in practice since in general only quantized CDI is available at the transmitter due to rate limitations on the feedback channel.

The main contribution of this paper is to study the impact of quantized CDI on the secrecy rate achievable by AN-assisted beamforming. Although the optimal signaling scheme is unknown for cases without knowledge of the eavesdropper's channel, AN-assisted beamforming has been shown to be optimal in the high SNR regime when the transmitter has full knowledge of the legitimate receiver's channel and is equipped with a large number of antennas [4]. This scheme is also optimal when the transmitter perfectly knows both the legitimate receiver's and the eavesdropper's channels [6]. When only quantized CDI is available at the transmitter, the AN that was originally intended to disrupt the eavesdropper's reception may now leak into the legitimate channel, causing degradation in the achievable secrecy rate. We refer to this as the *AN leakage problem*.

In this work, we first examine the optimal power allocation between the message-bearing signal and the AN for a given number of feedback bits B. We show that, when B is sufficiently large (good CDI quality), the power should be allocated evenly among the message-bearing signal and the AN; whereas, when B is small (poor CDI quality), the power allocated to the AN must be more conservative in order to limit the effects of AN leakage. Moreover, when B is fixed, we

Manuscript received March 12, 2010; revised September 16, 2010; accepted December 6, 2010. The associate editor coordinating the review of this paper and approving it for publication was G. Colavolpe.

The material in this paper was presented in part at the IEEE International Conference on Communications (ICC), South Africa, 2010, and at the IEEE International Symposium on Information Theory (ISIT), Korea, 2009.

The authors are with the Institute of Communications Engineering and Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013 (e-mail: {linsc, changth}@mx.nthu.edu.tw, {ywhong, cychi}@ee.nthu.edu.tw).

This work was supported in part by the National Science Council, Taiwan, under grants NSC-98-2219-E-007-004, NSC-98-2218-E-009-008-MY3, NSC-98-2219-E-007-005, and NSC-98-2219-E-007-003.

Digital Object Identifier 10.1109/TWC.2011.010411.100374

observe that in contrast to the perfect CDI case, the achievable secrecy rate will be upper-bounded by a constant regardless of the transmission power. Therefore, to maintain a constant secrecy rate loss (compared to the perfect CDI case), we show that B must scale logarithmically with the transmission power P. In this work, the channel quality information (CQI) is assumed to be unknown at the transmitter. However, we are able to show that, in the case of quantized CDI, additional CQI at the transmitter provides little performance gains when P is sufficiently large, which justifies our interest at the CDI only. The results of this work are first examined for the multiple-input single-output single-antenna eavesdropper (MISOSE) case, where the transmitter has multiple antennas and both the receiver and the eavesdropper have only a single antenna. The results are later extended to the multiple-input multiple-output multiple-antenna-eavesdropper (MIMOME) scenario, where both the receiver and the eavesdropper are assumed to have multiple antennas.

The effects of quantized channel feedback on the transceiver design have been studied in the literature for both single user and multiuser downlink systems (without eavesdroppers), e.g., in [7]–[9] and references therein. However, to the best of our knowledge, these issues have not been addressed before in the context of secret communications. Compared to our previous works [10], [11], we improve the bit scaling law for the MISOSE case in [10] by removing some approximation steps, and provide detailed proofs for the results in [11]. The results on the AN power allocation and the impact of CQI have not been presented before.

The rest of this paper is organized as follows. In Section II, we provide the system model and background on AN-assisted beamforming. In Section III, we examine the optimal power allocation between the message-bearing signal and the AN for the MISOSE case. The MISOSE feedback bit scaling law is provided in Section IV and discussions on the impact of CQI are provided in Section V. In Section VI, we extend the bit scaling law to the MIMOME scenario. Simulation results are provided in Section VII. Finally, the conclusion is drawn in Section VIII.

Notations: We denote $\mathbf{I}_n$ as the $n \times n$ identity matrix, and denote $|\mathbf{B}|$ and $\mathrm{Tr}(\mathbf{B})$ as the determinant and trace of matrix $\mathbf{B}$, respectively. $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_n)$ means that $\mathbf{x}$ is a complex Gaussian random vector with zero mean and covariance matrix $\sigma^2\mathbf{I}_n$, and $x \sim \beta(a,b)$ means that x is a Beta-distributed random variable with parameters (a,b). $\mathbf{E}[\cdot]$ stands for the statistical expectation of a random variable, $H(\cdot)$ stands for the entropy of a random variable (vector), and $I(x;y)$ represents the mutual information between random variables (vectors) x and y. Almost-sure convergence is denoted by $\overset{a.s.}{\to}$. The 2-norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$. The function $[x]^+$ represents $\max\{x, 0\}$. The $\log$ and $\ln$ functions are with base 2 and natural number e respectively. For an event A, the indicator function $\mathbf{1}_{\{A\}}$ is 1 if A occurs, and is 0 otherwise.

II. System Model and Background

Let us consider a wireless system that consists of a transmitter, a legitimate receiver, and an eavesdropper with M_t, M_r, and M_e antennas, respectively, as shown in Figure 1. The system model presented in this section is focused only on the MISOSE case where $M_t > M_e = M_r = 1$; while extensions to the MIMOME case (where $M_e \geq 1$ and $M_r \geq 1$) will be presented in Section VI. Let $\mathbf{x}[i] \in \mathbb{C}^{M_t \times 1}$ be the symbol vector transmitted in the ith time slot under the average power constraint $\mathbf{E}\left[\|\mathbf{x}[i]\|^2\right] \leq P$. The signals received at the receiver and the eavesdropper are

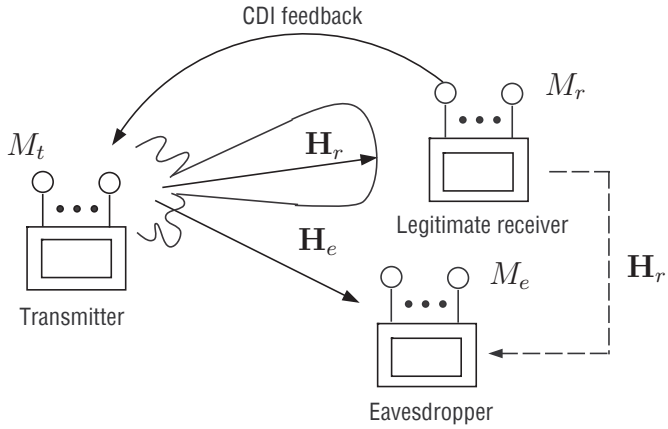

Fig. 1. A network diagram consisting of a multi-antenna transmitter, a legitimate receiver and an eavesdropper. Eavesdropper is assumed to know the full channel state information (including both CDI and channel quality information (CQI)) of the legitimate receiver.

$$y_r[i] = \mathbf{h}_r\mathbf{x}[i] + z_r[i] \quad \text{and} \quad y_e[i] = \mathbf{h}_e\mathbf{x}[i] + z_e[i], \tag{1}$$

respectively, where $\mathbf{h}_r, \mathbf{h}_e \in \mathbb{C}^{1\times M_t}$ are channel vectors of the receiver and the eavesdropper, respectively, with the same distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_t})$, and $z_r[i] \sim \mathcal{CN}(0,1)$, $z_e[i] \sim \mathcal{CN}(0, \sigma_e^2)$ are the additive white Gaussian noise (AWGN). Note that, in this case, the channel directions $\mathbf{g}_r = \mathbf{h}_r/\|\mathbf{h}_r\|$ and $\mathbf{g}_e = \mathbf{h}_e/\|\mathbf{h}_e\|$ are isotropically distributed on the unit sphere [12]. In this paper, we will assume that $\mathbf{h}_r$ and $\mathbf{h}_e$ are ergodic block-fading channels that remain constant over a sufficient amount of time for signal transmission and feedback, and the messages are coded across multiple fading blocks (c.f. (9) below). Our results can be readily extended to the case with ergodic fast-fading channels [3].

Most works on physical-layer secrecy consider the problem of reliably communicating a secret message from the transmitter to the legitimate receiver subject to a constraint on the information attainable by the eavesdropper. Consider a $(2^{nR}, n)$-code with an encoder $\mu_n : \mathcal{W}_n \to \mathbb{C}^{M_t\times n}$ that maps the message $w \in \mathcal{W}_n = \{1, 2, \ldots, 2^{nR}\}$ into a length-n codeword $\{\mathbf{x}[i]\}_{i=1}^n$ and a decoder ν_n at the legitimate receiver that maps the received sequence $\{y_r[i]\}_{i=1}^n$ to an estimated message $\hat{w} \in \mathcal{W}_n$. Define the error event as $\mathcal{E}_n = \{\hat{w} \neq w\}$. Perfect secrecy and secrecy capacity are defined as follows.

Definition 1 (Secrecy Capacity [2][4]): *Perfect secrecy is achievable with rate R if, for any $\varepsilon' > 0$, there exists a sequence of $(2^{nR}, n)$-codes and an integer n_0 such that, for any $n > n_0$,*

$$\Pr(\mathcal{E}_n) \leq \varepsilon', \text{ and } H(w|\{y_e[i]\}_{i=1}^n, \mathbf{h}_r^n, \mathbf{h}_e^n)/n > R - \varepsilon',$$

where w is the secret message, $\mathbf{h}_r^n$ and $\mathbf{h}_e^n$ are the collections of $\mathbf{h}_r$ and $\mathbf{h}_e$ over code length n, respectively. The secrecy capacity *is the supremum of all achievable secrecy rates.*

Note that the secrecy rate considered here is achieved by encoding over multiple channel states and the perfect secrecy constraint must be satisfied for all $n > n_0$. This implies that no secrecy outage [13] is allowed. In delay-limited applications, such perfect secrecy condition may not be achievable and, thus, a tradeoff exists between secrecy rate and secrecy outage probability. These issues have been discussed in [13] and are beyond the scope of this paper.

In this paper, we consider the case where the eavesdropper's channel $\mathbf{h}_e$ and noise variance σ_e^2 are unknown to the transmitter. In this case, the optimal signaling for the MISOSE channel is unknown (except for special cases given in [2] [4]). To guarantee secrecy, Goel and Negi proposed in [5] the use of *artificial noise* (AN) to disrupt the reception of eavesdropper while beamforming the message towards the legitimate receiver. The transmitted signal is given by

$$\mathbf{x}[i] = \mathbf{p}s[i] + \mathbf{Q}\mathbf{a}[i], \tag{2}$$

where $s[i]$ is the message-bearing signal with $\mathbf{E}[|s[i]|^2] = \sigma_s^2$, $\mathbf{p} \in \mathbb{C}^{M_t \times 1}$ is the normalized beamforming vector for $s[i]$, $\mathbf{Q} \in \mathbb{C}^{M_t \times (M_t-1)}$ is a matrix with columns that form an orthonormal basis for the AN subspace, and $\mathbf{a}[i]$ is the AN vector which is a random Gaussian vector with distribution $\mathcal{CN}(\mathbf{0}, \sigma_a^2 \mathbf{I}_{M_t-1})$. Here $s[i]$ and $\mathbf{a}[i]$ are assumed to be independent.

MISOSE Secrecy Rate with Quantized CDI

Following the studies on quantized channel feedback in [7], [8], we assume that the receiver is able to obtain perfect knowledge of $\mathbf{h}_r$, but can send back only a quantized version of the CDI, i.e., $\mathbf{g}_r = \mathbf{h}_r/\|\mathbf{h}_r\|$, to the transmitter due to limited feedback channel bandwidth. The CQI, i.e., $\|\mathbf{h}_r\|$, is assumed unknown at the transmitter. However, as we show later in Section V, lack of such information under the quantized feedback scenario has little impact on the secrecy rate when the transmission power P is large.

Suppose that the CDI $\mathbf{g}_r$ is quantized into one of 2^B unit-norm channel vectors in the codebook $\mathcal{C} \triangleq \{\mathbf{c}_1, ..., \mathbf{c}_{2^B}\}$ according to the minimum distance criterion [7], and the corresponding index, denoted by $\ell^\star = \arg\max_{\ell=1,...,2^B} \|\mathbf{g}_r \mathbf{c}_\ell^H\|$, is sent back to the transmitter. Let $\hat{\mathbf{g}}_r \triangleq \mathbf{c}_{\ell^\star}$ be the corresponding quantized CDI vector. The quantization cell associated with $\mathbf{c}_\ell \in \mathcal{C}$ is given by

$$\mathcal{V}_\ell = \{\mathbf{g}|\ |\mathbf{g}\mathbf{c}_\ell^H|^2 \geq |\mathbf{g}\mathbf{c}_j^H|^2\ \forall\ j \neq \ell\}. \tag{3}$$

To gain analytical insights on the impact of quantized CDI, we adopt the random vector quantization (RVQ) codebook [7], [8], where each codeword is a randomly and independently generated M_t-dimensional unit-norm complex Gaussian vector and, thus, is isotropically distributed in $\mathbb{C}^{1\times M_t}$ [12]. Moreover, we will also utilize the quantization cell approximation (QCA) model [7], [14], where each quantization cell $\mathcal{V}_\ell$ is approximated by a Voronoi region of a spherical cap with the surface area approximately equal to 2^{-B} of the total surface area of the M_t-dimensional unit sphere. It has been shown in [7], [8] that the behavior of RVQ can be closely approximated by this model, even for small B. Specifically, the quantization cell $\mathcal{V}_\ell$ in (3) is approximated by

$$\mathcal{V}_\ell \approx \{\mathbf{g}|\ |\mathbf{g}\mathbf{c}_\ell^H|^2 \geq 1-\delta\}, \quad \delta = 2^{-\frac{B}{M_t-1}}. \tag{4}$$

Define $|\mathbf{g}_r \hat{\mathbf{g}}_r^H|^2 = |\mathbf{g}_r \mathbf{c}_{\ell^\star}^H|^2 \triangleq \cos^2\theta \geq 1-\delta$. The cumulative distribution function (CDF) of $\sin^2\theta$ is [7]

$$\Pr(\sin^2\theta \leq x) = \begin{cases} 2^B x^{M_t-1}, & \text{for } 0 \leq x \leq \delta \\ 1, & \text{otherwise.} \end{cases} \tag{5}$$

When only the quantized CDI $\hat{\mathbf{g}}_r$ is available at the transmitter, the beamforming vector $\mathbf{p}$ and the matrix $\mathbf{Q}$ in (2) are set to $\hat{\mathbf{g}}_r^H$ and $\mathbf{N}_{\hat{\mathbf{g}}_r}$, respectively, where $\mathbf{N}_{\hat{\mathbf{g}}_r}$ has columns that form an orthonormal basis for the null space of $\hat{\mathbf{g}}_r$. Therefore, the transmitted signal is given by

$$\mathbf{x}[i] = \hat{\mathbf{g}}_r^H s[i] + \mathbf{N}_{\hat{\mathbf{g}}_r}\mathbf{a}[i]. \tag{6}$$

Note that finding the optimal $\mathbf{p}$ and $\mathbf{Q}$ under quantized CDI is in general difficult due to AN leakage. Our choice of using $\mathbf{p} = \hat{\mathbf{g}}_r$ is motivated by [7], [9] where the multi-user interference in their case has a similar effect as the AN leakage in our case. To satisfy the power constraint $\mathbf{E}[\|\mathbf{x}[i]\|^2] = \sigma_s^2 + (M_t-1)\sigma_a^2 \leq P$, we set $\sigma_s^2 = \alpha P$ and $\sigma_a^2 = \frac{(1-\alpha)P}{M_t-1}$, where $\alpha \in [0,1]$ denotes the fraction of power allocation. Since the eavesdropper's noise variance σ_e^2 is assumed unknown to the transmitter, we shall consider throughout the rest of this paper the worst-case scenario where $\sigma_e^2 = 0$. By (1), the received signals of the receiver and the eavesdropper are given by

$$\hat{y}_r[i] = \|\mathbf{h}_r\|(\mathbf{g}_r\hat{\mathbf{g}}_r^H)s[i] + \|\mathbf{h}_r\|(\mathbf{g}_r\mathbf{N}_{\hat{\mathbf{g}}_r})\mathbf{a}[i] + z_r[i], \tag{7}$$

and

$$\hat{y}_e[i] = \mathbf{h}_e\hat{\mathbf{g}}_r^H s[i] + \mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\mathbf{a}[i], \tag{8}$$

respectively. One can observe from (7) that, due to imperfect CDI, the AN that was originally intended for the eavesdropper also interferes with the legitimate receiver. By assuming that $s[i]$ is Gaussian and from (7) and (8), the achievable secrecy rate under quantized CDI is

$$\begin{aligned}\hat{R}(\alpha) &= (I(s;\hat{y}_r|\mathbf{h}_r) - I(s;\hat{y}_e|\mathbf{h}_r,\mathbf{h}_e))^+ \\ &= \left(\mathbf{E}\left[\log\left(1+\frac{\|\mathbf{h}_r\|^2\cos^2\theta\cdot\alpha P}{\|\mathbf{h}_r\|^2\sin^2\theta\left(\frac{1-\alpha}{M_t-1}\right)P+1}\right)\right]\right. \\ &\quad \left. -\mathbf{E}\left[\log\left(1+\frac{\left|\mathbf{h}_e\hat{\mathbf{g}}_r^H\right|^2\alpha}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2\left(\frac{1-\alpha}{M_t-1}\right)}\right)\right]\right)^+, \end{aligned} \tag{9}$$

where we have used the fact that $\|\mathbf{g}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2 = 1-|\mathbf{g}_r\hat{\mathbf{g}}_r^H|^2 = \sin^2\theta$. The secrecy rate in (9) is achievable by coding across multiple fading states, as shown in [1], [2][1]. However, it is easy

[1]Note that, without CQI, the variable-rate coding in [2] can not be applied and, thus, the result in (9) is derived based on the constant-rate coding scheme, which can also be found in [2].

to see that, as P goes to infinity, this secrecy rate converges to the constant

$$\left(\mathbf{E}\left[\log\left(1+\frac{\alpha\cos^2\theta}{\sin^2\theta\cdot(1-\alpha)/(M_t-1)}\right)-\log\left(1+\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2\alpha}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2(1-\alpha)/(M_t-1)}\right)\right]\right)^+. \quad (10)$$

In fact, this secrecy rate is bounded even for arbitrary distributions of $s[i]$ and with the CQI at the transmitter, as we will show later in Section V.

In contrast, in the perfect CDI case where $\mathbf{g}_r$ is perfectly known at the transmitter, one can impose AN perfectly in the null space of $\mathbf{g}_r$ (i.e., choosing $\mathbf{Q}=\mathbf{N}_{\mathbf{g}_r}$) so that there is no noise leakage, that is, $\mathbf{g}_r\mathbf{Q}=\mathbf{g}_r\mathbf{N}_{\mathbf{g}_r}=\mathbf{0}$ (or $\cos\theta=1$). The resultant secrecy rate in (9) is given by

$$R(\alpha)=\left(\mathbf{E}\left[\log(1+\|\mathbf{h}_r\|^2\alpha P)-\log\left(1+\frac{|\mathbf{h}_e\mathbf{g}_r^H|^2\alpha}{\|\mathbf{h}_e\mathbf{N}_{\mathbf{g}_r}\|^2(1-\alpha)/(M_t-1)}\right)\right]\right)^+, \quad (11)$$

which obviously can be made arbitrarily large by increasing P.

Notice that, in the worst-case scenario where $\sigma_e^2=0$, AN is essential to achieving a nonzero secrecy rate. Specifically, if one sets α equal to 1 (no AN), then the second terms inside $(\cdot)^+$ in (9) and (11) will go to infinity, leading to a zero secrecy rate. However, AN may also cause significant loss in secrecy rate under imperfect CDI due to AN leakage. Therefore, the power allocation between the message-bearing signal $s[i]$ and the AN $\mathbf{a}[i]$ must be carefully determined when the number of feedback bits B is limited.

III. Power Allocation of Signal and Artificial Noise under Quantized CDI

In this section, we study the power allocation between message-bearing signal $s[i]$ and AN $\mathbf{a}[i]$ for a given number of feedback bits B. To this end, we first present a useful lemma which shows that for M_t large, the second term inside $(\cdot)^+$ in (9) will converge to a fixed value.

Lemma 1: *Let $\mathbf{E}_1(x)=\int_x^\infty\frac{e^{-t}}{t}dt$ be the exponential integral, it follows that*

$$\lim_{M_t\to\infty}\mathbf{E}\left[\log\left(1+\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}\frac{\alpha}{(1-\alpha)/(M_t-1)}\right)\right]=\frac{1}{\ln 2}\mathbf{E}_1\left(\frac{1-\alpha}{\alpha}\right)\exp\left(\frac{1-\alpha}{\alpha}\right). \quad (12)$$

The proof is given in Appendix A. In this case, for M_t sufficiently large, the achievable secrecy rate in (9) can be expressed approximately as

$$\hat{R}(\alpha)\cong\left(\mathbf{E}\left[\log\left(\frac{\gamma_{\hat{\mathbf{g}}_r}\alpha P}{\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha)P+1}\right)\right]-\frac{1}{\ln 2}\mathbf{E}_1\left(\frac{1-\alpha}{\alpha}\right)\exp\left(\frac{1-\alpha}{\alpha}\right)\right)^+ \quad (13)$$

where

$$\gamma_{\hat{\mathbf{g}}_r}\triangleq\|\mathbf{h}_r\|^2\cos^2\theta \ \text{ and } \ \gamma_{\hat{\mathbf{g}}_r^\perp}\triangleq\frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta \quad (14)$$

are the squared channel gains of $\mathbf{h}_r$ that fall, respectively, in the direction of $\hat{\mathbf{g}}_r$ and in the orthogonal subspace of $\hat{\mathbf{g}}_r$ (while normalized by their respective dimensions). Let us define

$$F(\alpha)\triangleq\mathbf{E}\left[\log\left(\frac{\gamma_{\hat{\mathbf{g}}_r}\alpha P}{\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha)P+1}\right)\right]-\frac{1}{\ln 2}\mathbf{E}_1\left(\frac{1-\alpha}{\alpha}\right)\exp\left(\frac{1-\alpha}{\alpha}\right)$$

such that $\hat{R}(\alpha)\cong(F(\alpha))^+$. By taking the derivative of $F(\alpha)$ and setting it to zero, i.e., $\frac{\partial F(\alpha)}{\partial\alpha}=0$, it follows that the optimal α, denoted by $\alpha^\star$, must satisfy the necessary condition

$$\mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r}P(\gamma_{\hat{\mathbf{g}}_r^\perp}P+1)(1-\alpha^\star)}{(\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P+1)(\gamma_{\hat{\mathbf{g}}_r}\alpha^\star P+\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P+1)}\right]=\frac{1}{\alpha^\star}-\frac{1-\alpha^\star}{(\alpha^\star)^2}\mathbf{E}_1\left(\frac{1-\alpha^\star}{\alpha^\star}\right)\exp\left(\frac{1-\alpha^\star}{\alpha^\star}\right). \quad (15)$$

Although a closed-form solution of $\alpha^\star$ is not easily computable due to the intractability of the expectation term in (15), explicit results can be obtained for two interesting special cases, namely, the case of *weak AN leakage* and the case of *strong AN leakage*. In the case of weak AN leakage, B is assumed sufficiently large such that AN leakage is much smaller than AWGN, i.e. $\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha)P\ll 1$. On the other hand, in the case of strong AN leakage, B is assumed to be small such that $\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha)P\gg 1$. For the former case, we have the following proposition:

Proposition 1 (Weak AN Leakage): *Let α_w be the value that satisfies $\frac{1-\alpha_w}{\alpha_w}=d(\alpha_w)$, where*

$$d(\alpha^\star)\triangleq\frac{1}{\alpha^\star}-\frac{1-\alpha^\star}{(\alpha^\star)^2}\mathbf{E}_1\left(\frac{1-\alpha^\star}{\alpha^\star}\right)\exp\left(\frac{1-\alpha^\star}{\alpha^\star}\right).$$

Then, for $B>(M_t-1)\log(P/\epsilon)$ and any ϵ such that $\frac{1-\alpha_w}{\alpha_w}>\epsilon>0$, the optimal power allocation fraction $\alpha^\star$ of $\hat{R}(\alpha)$ in (13) satisfies $\alpha_w-\epsilon_l<\alpha^\star<\alpha_w+\epsilon_u$ for M_t sufficiently large, where both $\epsilon_u,\epsilon_l>0$ and will approach 0 as ϵ approaches 0.

The proof is provided in Appendix B. Note that $d(\alpha^\star)$ comes from (15). Numerically, we can show that $\alpha_w\approx 0.554$. The fact that $\alpha^\star$ is close to one half can also be roughly observed from (9). Specifically, with B sufficiently large, AN leakage will be negligible and the first term in (9) will be proportional to $\log\alpha$ when M_t is large. Also, by applying the approximation $\mathbf{E}_1(x)e^x\approx\ln(\frac{1+x}{x})$ (see [15]) and the results of Lemma 1, the second term becomes $-\log(1-\alpha)$. In this

case, (9) will be approximately proportional to $\log(\alpha(1-\alpha))$ and the optimal $\alpha^\star$ will be close to 0.5. The results of Proposition 1 shows that, when channel knowledge is sufficiently accurate, one should spend almost half the power on AN to disrupt the eavesdroppers reception. However, this is not the case under strong AN leakage as shown in the following proposition.

Proposition 2 (Strong AN Leakage): *For any $\epsilon > 0$ and $P \geq 1/(\epsilon\delta/2)$, the optimal power allocation fraction $\alpha^\star$ of $\hat{R}(\alpha)$ in* (13) *satisfies*

$$\alpha^\star > \min\left\{1 - \frac{\epsilon}{1+\epsilon}, 1 - \frac{1}{\epsilon(\delta/2)P}\right\}, \tag{16}$$

for $\eta \triangleq B/M_t > 0$ fixed and for M_t sufficiently large, where δ was defined in (4).

The proof is given in Appendix C. According to Proposition 2, when $P \gg 1/(\epsilon\delta/2) = (2/\epsilon)2^{B/(M_t-1)}$ [by (4)], or equivalently when $B \ll (M_t-1)\log(\epsilon P/2)$, the optimal $\alpha^\star$ approaches 1 since ϵ can be chosen arbitrarily small. This implies that, when the channel quantization is coarse, i.e., B is not large enough, one should allocate less power for the AN simply owing to the severe AN leakage. While we have examined the optimal power allocation strategy under quantized CDI, this is not yet sufficient to guarantee a bounded secrecy rate loss compared to the perfect CDI case, as we will present in the next section.

IV. Scaling of the Number of Feedback Bits in the MISOSE case

In the previous section, the optimal power allocation between the message-bearing signal and the AN has been presented to reduce the secrecy rate loss under a given transmission power P. However, as P increases, the secrecy rate loss compared to the perfect CDI case will become arbitrarily large for a fixed B since the achievable secrecy rate under perfect CDI (c.f. (11)) increases without bound while that under quantized CDI (cf. (9)) is bounded for any value of α. To resolve this problem, we derive in this section the value of B that is needed for maintaining a constant secrecy rate loss.

Let us define the secrecy rate loss as

$$\Delta R \triangleq \max_\alpha R(\alpha) - \max_\alpha \hat{R}(\alpha) \leq R(\alpha_{\rm p}^\star) - \hat{R}(\alpha_{\rm p}^\star) \tag{17}$$

where $\alpha_{\rm p}^\star = \arg\max_\alpha R(\alpha)$. The bound follows since the secrecy rate under quantized CDI is generally not maximized with $\alpha = \alpha_{\rm p}^\star$. Note, by Proposition 1, that $\alpha_{\rm p}^\star \approx 0.5$ when M_t is large. Recalling (9) and (11), and the worst-case assumption of $\sigma_e^2 = 0$, we have

$$R(\alpha_{\rm p}^\star) = \left(\mathbf{E}\left[\log(1+\|\mathbf{h}_r\|^2\alpha_{\rm p}^\star P)\right] - \mathbf{E}\left[\log\left(1+\frac{|\mathbf{h}_e\mathbf{g}_r^H|^2\alpha_{\rm p}^\star}{\|\mathbf{h}_e\mathbf{N}_{\mathbf{g}_r}\|^2(1-\alpha_{\rm p}^\star)/(M_t-1)}\right)\right]\right)^+ \tag{18}$$

and

$$\hat{R}(\alpha_{\rm p}^\star) = \left(\mathbf{E}\left[\log\left(1+\frac{\|\mathbf{h_r}\|^2\cos^2\theta\cdot\alpha_{\rm p}^\star P}{\|\mathbf{h}_r\|^2\sin^2\theta\left(\frac{1-\alpha_{\rm p}^\star}{M_t-1}\right)P+1}\right)\right] - \mathbf{E}\left[\log\left(1+\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2\alpha_{\rm p}^\star}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2\left(\frac{1-\alpha_{\rm p}^\star}{M_t-1}\right)}\right)\right]\right)^+. \tag{19}$$

As noted in [7], for any two independent and isotropically distributed unit-norm vectors $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{C}^{1\times M_t}$, their squared inner product will be Beta-distributed with parameters $(1, M_t-1)$, i.e., $|\mathbf{w}_1\mathbf{w}_2^H|^2 \sim \beta(1, M_t-1)$. Applying this result to $\mathbf{g}_e = \mathbf{h}_e/\|\mathbf{h}_e\|$ and $\mathbf{g}_r$, one can have $|\mathbf{h}_e\mathbf{g}_r^H|^2 = \|\mathbf{h}_e\|^2\beta(1, M_t-1)$. This implies that the second term inside $(\cdot)^+$ in (18) can be written as

$$\begin{aligned}&\mathbf{E}\left[\log\left(1+\frac{|\mathbf{h}_e\mathbf{g}_r^H|^2\alpha_{\rm p}^\star}{\|\mathbf{h}_e\mathbf{N}_{\mathbf{g}_r}\|^2\left(\frac{1-\alpha_{\rm p}^\star}{M_t-1}\right)}\right)\right]\\ =&\mathbf{E}\left[\log\left(1+\frac{\beta(1,M_t-1)\alpha_{\rm p}^\star}{(1-\beta(1,M_t-1))\left(\frac{1-\alpha_{\rm p}^\star}{M_t-1}\right)}\right)\right],\end{aligned}$$

where we have used the fact that $\|\mathbf{h}_e\mathbf{N}_{\mathbf{g}_r}\|^2 = \|\mathbf{h}_e\|^2 - |\mathbf{h}_e\mathbf{g}_r^H|^2$. The same argument also applies to the second term inside $(\cdot)^+$ in (19) since $\mathbf{g}_e$ and $\hat{\mathbf{g}}_r$ are also independent and isotropically distributed under RVQ. Therefore, the second terms inside $(\cdot)^+$ in both (18) and (19) yield the same value, and thus the secrecy rate loss in (17) can be bounded as

$$\begin{aligned}\Delta R \leq\ &\mathbf{E}[\log(1+\|\mathbf{h}_r\|^2\cdot\alpha_{\rm p}^\star P)]\\ &-\mathbf{E}\left[\log\left(1+\frac{\|\mathbf{h}_r\|^2\cos^2\theta\cdot\alpha_{\rm p}^\star P}{\|\mathbf{h}_r\|^2\sin^2\theta\cdot\frac{(1-\alpha_{\rm p}^\star)}{(M_t-1)}P+1}\right)\right].\end{aligned} \tag{20}$$

By further evaluating the expectation and by applying Jensen's inequality [16], we show in Appendix D the following theorem:

Theorem 1: *The secrecy rate loss between perfect and quantized CDI is upper-bounded by*

$$\begin{aligned}\Delta R <&\log\left[\frac{M_t(1-\alpha_{\rm p}^\star)P2^{\frac{-B}{M_t-1}}+(M_t-1)}{(M_t-1)(1-2^{\frac{B}{M_t-1}})}\right]\\ &+\log\left[\left(1+\frac{1}{(M_t-1)\alpha_{\rm p}^\star P}\right)\right].\end{aligned} \tag{21}$$

This theorem implies that, to maintain a constant secrecy rate loss of c, i.e., $\Delta R \leq c$, B must be chosen such that

$$\begin{aligned}B \geq (M_t-1)\log\Bigg[&\frac{M_t}{M_t-1}\cdot\frac{(1-\alpha_{\rm p}^\star)P}{2^{c-c'(P,M_t)}-1}\\ &+\frac{2^{c-c'(P,M_t)}}{2^{c-c'(P,M_t)}-1}\Bigg],\end{aligned} \tag{22}$$

where

$$c'(P,M_t) = \log\left((M_t-1)\alpha_{\rm p}^\star P+1\right) - \log\left((M_t-1)\alpha_{\rm p}^\star P\right).$$

Notice that $c'(P, M_t)$ approaches to 0 as M_t or P increases. Hence, to guarantee a constant secrecy rate loss, B must scale linearly with M_t or logarithmically with P, i.e., $B = \Omega(M_t \log P)$ in big-Ω notation [17].

V. Discussions on the Impact of CQI and Input Distributions

In the previous sections, we have shown that the achievable secrecy rate under quantized CDI is bounded under Gaussian $s[i]$ and in the absence of CQI at the transmitter. This result has motivated our study on the scaling of B in order to maintain a constant secrecy rate loss. In this section, we present a stronger result: The distribution of $s[i]$ and knowledge of CQI in fact have limited impact on the achievable secrecy rate, and therefore the scaling of B in (22) is essential to the secrecy rate loss control. Specifically, we prove in Appendix E the following theorem:

Theorem 2: *Consider the MISOSE signal model in* (7) *and* (8) *under quantized CDI. Suppose that the message-bearing signal $s[i]$ follows an arbitrary statistical distribution, and that the transmitter perfectly knows the CQI $\|\mathbf{h}_r\|$. Then the achievable secrecy rate $\hat{R}$ is upper bounded by a constant $\hat{R}_{UB}$ as*

$$\hat{R} \le \hat{R}_{UB} \triangleq \mathbf{E}\left[\left(\log \frac{|\mathbf{g}_r \hat{\mathbf{g}}_r^H|^2}{\|\mathbf{g}_r \mathbf{N}_{\hat{\mathbf{g}}_r}\|^2} \frac{\|\mathbf{g}_e \mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}{|\mathbf{g}_e \hat{\mathbf{g}}_r^H|^2}\right)^+\right], \quad (23)$$

where $\mathbf{g}_e = \mathbf{h}_e/\|\mathbf{h}_e\|$.

It can be observed that this bound neither depends on the power nor on the channel gains, but depends only on the channel directions and B (since both $\hat{\mathbf{g}}_r$ and $\mathbf{N}_{\hat{\mathbf{g}}_r}$ are functions of B). More interestingly, for M_t sufficiently large, the difference between this bound and the achievable secrecy rate in (9) can be small as $P \to \infty$. This is stated in the following corollary.

Corollary 1: Given the values of $\alpha \in (0, 1)$ and $\eta \triangleq B/M_t > 0$, for any $\varepsilon > 0$, the difference between the upper bound in (23) and the achievable secrecy rate in (9) is no greater than $|\log \alpha| + 0.83 + \varepsilon$, when M_t is sufficiently large and P is sufficiently larger than $(2/(1-\alpha))2^{\frac{B}{M_t - 1}}$.

The proof of Corollary 1 is given in Appendix F. It is worthwhile to note that, with α fixed, the upper bound given in (23) can increase without bound by increasing the ratio B/M_t since $\|\mathbf{g}_r \mathbf{N}_{\hat{\mathbf{g}}_r}\|$ will go to zero in this case. Therefore, by Corollary 1, the ratio between the achievable secrecy rate in (9) and the upper bound in (23) will approach 1, when α is fixed and when M_t and P are sufficiently large. That is, the secrecy rate loss will become negligible compared to the achievable secrecy rate. Since the bound in (23) holds for arbitrary distributions of $s[i]$ and for full CQI at the transmitter, Corollary 1 further implies that Gaussian $s[i]$ is nearly optimal and the CQI at the transmitter does not provide much gain for the achievable secrecy rate in (9).

VI. Extensions to the MIMOME Scenario

In this section, we extend our studies to the MIMOME scenario where M_r, $M_e \ge 1$, and derive the scaling of the number of feedback bits B under this scenario. The received signals at the receiver and the eavesdropper can be expressed as

$$\mathbf{y}_r = \mathbf{H}_r \mathbf{x} + \mathbf{z}_r \quad \text{and} \quad \mathbf{y}_e = \mathbf{H}_e \mathbf{x}, \quad (24)$$

where we have assumed that the eavesdropper does not suffer from noise. As in the MISOSE case, the channel matrices $\mathbf{H}_r \in \mathbb{C}^{M_r \times M_t}$ and $\mathbf{H}_e \in \mathbb{C}^{M_e \times M_t}$ are assumed to be ergodic block faded, with each entry being i.i.d. complex Gaussian with zero mean and unit variance.

Assume that $M_t \ge M_r + M_e$. Let $\mathbf{H}_r = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^H$ be the singular value decomposition of $\mathbf{H}_r$, where $\mathbf{V} \in \mathbb{C}^{M_r \times M_r}$ is a unitary matrix, $\mathbf{\Sigma} \in \mathbb{C}^{M_r \times M_r}$ is a diagonal matrix with the singular values of $\mathbf{H}_r$ being the diagonal elements, and $\mathbf{U} \in \mathbb{C}^{M_t \times M_r}$ is a semi-unitary matrix. The transmitter knows a quantized version of the CDI, i.e., $\hat{\mathbf{U}} \in \mathbb{C}^{M_t \times M_r}$. Let $\mathbf{N}_{\hat{\mathbf{U}}} \in \mathbb{C}^{M_t \times (M_t - M_r)}$ be a matrix whose columns form an orthonormal basis for the null space of $\hat{\mathbf{U}}^H$, the transmitted signal is given by $\mathbf{x} = \hat{\mathbf{U}}\mathbf{s} + \mathbf{N}_{\hat{\mathbf{U}}}\mathbf{a}$, where $\mathbf{s} \in \mathbb{C}^{M_r}$ is the message-bearing signal with distribution $\mathcal{CN}(0, \sigma_s^2 \mathbf{I}_{M_r})$, and $\mathbf{a} \in \mathbb{C}^{M_t - M_r}$ is the imposed AN with distribution $\mathcal{CN}(0, \sigma_a^2 \mathbf{I}_{M_t - M_r})$. We let $\sigma_s^2 = \alpha P/M_r$ and $\sigma_a^2 = (1-\alpha)P/(M_t - M_r)$ where $0 \le \alpha \le 1$, in order to constrain the transmission power within P. Substituting $\mathbf{x} = \hat{\mathbf{U}}\mathbf{s} + \mathbf{N}_{\hat{\mathbf{U}}}\mathbf{a}$ and $\mathbf{H}_r = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^H$ into (24) yields

$$\mathbf{y}_r = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^H\hat{\mathbf{U}}\mathbf{s} + \mathbf{V}\mathbf{\Sigma}\mathbf{U}^H\mathbf{N}_{\hat{\mathbf{U}}}\mathbf{a} + \mathbf{z}_r \quad (25)$$

$$\mathbf{y}_e = \mathbf{H}_e\hat{\mathbf{U}}\mathbf{s} + \mathbf{H}_e\mathbf{N}_{\hat{\mathbf{U}}}\mathbf{a}. \quad (26)$$

Since $\mathbf{U}^H\mathbf{N}_{\hat{\mathbf{U}}} \ne \mathbf{0}$, the legitimate receiver will experience AN leakage, resulting in the achievable secrecy rate (27) given in the top of the next page.

Similar to the MISOSE case, the secrecy rate due to imperfect CDI also converges to a constant as P goes to infinity. For the case where perfect CDI is available at the transmitter, the transmit signal is given by $\mathbf{x} = \mathbf{U}\mathbf{s} + \mathbf{N}_{\mathbf{U}}\mathbf{a}$. Since $\mathbf{U}^H\mathbf{U} = \mathbf{I}_{M_r}$ and $\mathbf{U}^H\mathbf{N}_{\mathbf{U}} = \mathbf{0}$, the achievable MIMOME secrecy rate with perfect CDI [5] is given by

$$R_{\mathrm{M}}(\alpha) = \Bigg(\mathbf{E}\left[\log\left|\mathbf{I}_{M_r} + \mathbf{\Sigma}^2\sigma_s^2\right|\right] - \mathbf{E}\left[\log\left|\mathbf{I}_{M_r} + \sigma_s^2\mathbf{U}^H\mathbf{H}_e^H\left(\sigma_a^2\mathbf{H}_e\mathbf{N}_{\mathbf{U}}\mathbf{N}_{\mathbf{U}}^H\mathbf{H}_e^H\right)^{-1}\mathbf{H}_e\mathbf{U}\right|\right]\Bigg)^+. \quad (28)$$

The MIMOME secrecy rate in (28) can be made arbitrarily large by increasing P since the first term inside $(\cdot)^+$ increases without bound while the second term inside $(\cdot)^+$ is bounded as $P \to \infty$. Therefore, for a given B, the secrecy rate loss compared to the perfect CDI case will become unbounded as P increases.

A. Random Quantization Codebook Model

To analyze the achievable secrecy rate under quantized CDI, let us consider the random quantization codebook $\mathcal{C} = \{\mathbf{C}_1, \ldots, \mathbf{C}_{2^B}\}$, where the 2^B semi-unitary matrices

$$\hat{R}_M(\alpha) = \left(\mathbf{E}\left[\log\left| \mathbf{I}_{M_r} + \sigma_s^2 \hat{\mathbf{U}}^H \mathbf{U}\mathbf{\Sigma} \left(\mathbf{I}_{M_r} + \sigma_a^2 \mathbf{\Sigma}\mathbf{U}^H \mathbf{N}_{\hat{\mathbf{U}}} \mathbf{N}_{\hat{\mathbf{U}}}^H \mathbf{U}\mathbf{\Sigma} \right)^{-1} \mathbf{\Sigma}\mathbf{U}^H \hat{\mathbf{U}} \right| \right] - \mathbf{E}\left[\log\left| \mathbf{I}_{M_r} + \sigma_s^2 \hat{\mathbf{U}}^H \mathbf{H}_e^H \left(\sigma_a^2 \mathbf{H}_e \mathbf{N}_{\hat{\mathbf{U}}} \mathbf{N}_{\hat{\mathbf{U}}}^H \mathbf{H}_e^H \right)^{-1} \mathbf{H}_e \hat{\mathbf{U}} \right| \right] \right)^+ . \quad (27)$$

are chosen independently and isotropically over the $M_t \times M_r$ Grassmann manifold (which is the set of all M_r-dimensional subspaces in an M_t-dimensional space) [18]. The quantized CDI must meet

$$\hat{\mathbf{U}} = \arg \min_{\mathbf{C} \in \mathcal{C}} d^2(\mathbf{U}, \mathbf{C}), \quad (29)$$

where $d(\mathbf{U}, \mathbf{C}) = M_r - \mathrm{Tr}(\mathbf{U}^H \mathbf{C}\mathbf{C}^H \mathbf{U})$ is the chordal distance between $\mathbf{U}$ and $\mathbf{C}$ [18]. In this model, the average distortion between $\hat{\mathbf{U}}$ and $\mathbf{U}$ can be upper bounded as [18]

$$D \triangleq \mathbf{E}[d^2(\mathbf{U}, \hat{\mathbf{U}})] \leq \frac{1}{m} \Gamma\left(\frac{1}{m}\right) \Phi^{-1/m} \, 2^{-B/m}, \quad (30)$$

where $\Gamma(\cdot)$ is the Gamma function, $m = M_r(M_t - M_r)$, and $\Phi = \frac{1}{\Gamma(m+1)} \prod_{i=1}^{M_c} \frac{M_t - i + 1}{M_c - i + 1}$ where $M_c \triangleq M_r - 2[M_r - M_t/2]^+$. The following lemma describes the relation between $\mathbf{U}$ and $\hat{\mathbf{U}}$.

Lemma 2 ([9]): *Under the random quantization codebook model, the quantized CDI $\hat{\mathbf{U}}$ and the true CDI $\mathbf{U}$ satisfy*

$$\mathbf{U} = \hat{\mathbf{U}}\mathbf{X}\mathbf{Y} + \hat{\mathbf{S}}\mathbf{R}, \quad (31)$$

where $\mathbf{X} \in \mathbb{C}^{M_r \times M_r}$ is a unitary matrix and $\mathbf{Y} \in \mathbb{C}^{M_r \times M_r}$ is an upper triangular matrix which is statistically independent of $\mathbf{X}$ and satisfies $\mathbf{U}^H \hat{\mathbf{U}} \hat{\mathbf{U}}^H \mathbf{U} = \mathbf{Y}^H \mathbf{Y}$ and $\mathbf{E}\left(\mathbf{Y}^H \mathbf{Y}\right) = (1 - D/M_r)\, \mathbf{I}_{M_r}$. Moreover, the columns of $\hat{\mathbf{S}} \in \mathbb{C}^{M_t \times N}$ form an orthonormal basis for an isotropically distributed N-dimensional subspace in the range space of $\mathbf{N}_{\hat{\mathbf{U}}}$, where $N = \min\{M_t - M_r, M_r\}$, and $\mathbf{R} \in \mathbb{C}^{N \times M_r}$ satisfies $\mathbf{U}^H \mathbf{N}_{\hat{\mathbf{U}}} \mathbf{N}_{\hat{\mathbf{U}}}^H \mathbf{U} = \mathbf{R}^H \mathbf{R}$ and $\mathbf{E}\left(\mathbf{R}^H \mathbf{R}\right) = (D/M_r) \mathbf{I}_{M_r}$.

B. Scaling of the Number of Feedback Bits B

Given properties of the random quantization codebook in Lemma 2, we are ready to analyze the MIMOME secrecy rate loss. Since the elements of $\mathbf{H}_r$ are complex i.i.d. Gaussian, the channel direction $\mathbf{U}$ is isotropically distributed in the M_t-by-M_r Grassmann manifold, and thus $\mathbf{N}_{\mathbf{U}}$ is isotropically distributed in the M_t-by-$(M_t - M_r)$ Grassmann manifold. In fact, the quantized CDI $\hat{\mathbf{U}}$ has the same distribution as $\mathbf{U}$ under random quantization. Hence, by following similar arguments as in the MISOSE case and by the fact that $\mathbf{N}_{\mathbf{U}} \mathbf{N}_{\mathbf{U}}^H = \mathbf{I}_{M_r} - \mathbf{U}\mathbf{U}^H$ and $\mathbf{N}_{\hat{\mathbf{U}}} \mathbf{N}_{\hat{\mathbf{U}}}^H = \mathbf{I}_{M_r} - \hat{\mathbf{U}}\hat{\mathbf{U}}^H$, the second terms inside $(.)^+$ of (28) and (27) turn out to be identical. Hence the MIMOME secrecy rate loss $\Delta R_M \triangleq \max_\alpha R_M(\alpha) - \max_\alpha \hat{R}_M(\alpha)$ can be upper bounded as

$$\begin{aligned} \Delta R_M &\leq \mathbf{E}\left(\log\left| \mathbf{I}_{M_r} + \mathbf{\Sigma}^2 \sigma_s^{\star 2} \right| \right) - \mathbf{E}\left(\log\left| \mathbf{I}_{M_r} + \sigma_s^{\star 2}\, \hat{\mathbf{U}}^H \mathbf{U}\mathbf{\Sigma}\mathbf{\Omega}^{-1}\mathbf{\Sigma}\mathbf{U}^H \hat{\mathbf{U}} \right| \right) \\ &\leq \mathbf{E}\left(\log\left| \mathbf{I}_{M_r} + \mathbf{\Sigma}^2 \sigma_s^{\star 2} \right| \right) - \mathbf{E}\left(\log\left| \sigma_s^{\star 2}\, \hat{\mathbf{U}}^H \mathbf{U}\mathbf{\Sigma}\mathbf{\Omega}^{-1}\mathbf{\Sigma}\mathbf{U}^H \hat{\mathbf{U}} \right| \right), \end{aligned} \quad (32)$$

where $\mathbf{\Omega} \triangleq \mathbf{I}_{M_r} + \sigma_a^{\star 2} \mathbf{\Sigma}\mathbf{U}^H \mathbf{N}_{\hat{\mathbf{U}}} \mathbf{N}_{\hat{\mathbf{U}}}^H \mathbf{U}\mathbf{\Sigma}$, and $\sigma_s^{\star 2} = \alpha_{\mathrm{p},M}^\star P/M_r$, $\sigma_a^{\star 2} = (1 - \alpha_{\mathrm{p},M}^\star)P/(M_t - M_r)$ in which $\alpha_{\mathrm{p},M}^\star = \arg\max_\alpha R_M(\alpha)$. By Lemma 2, we have that $\hat{\mathbf{U}}^H \mathbf{U} = \hat{\mathbf{U}}^H \hat{\mathbf{U}} \mathbf{X}\mathbf{Y} + \hat{\mathbf{U}}^H \hat{\mathbf{S}}\mathbf{R} = \mathbf{X}\mathbf{Y}$. Since $\mathbf{X}$ and $\mathbf{Y}$ are statistically independent (by Lemma 2), it follows with probability one that $\hat{\mathbf{U}}^H \mathbf{U}$ is full rank [19]. Hence, the secrecy rate loss ΔR_M can be further bounded by

$$\begin{aligned} \Delta R_M &\leq \mathbf{E}\left[\log\left|\mathbf{I}_{M_r} + \mathbf{\Sigma}^{-2}/\sigma_s^{\star 2}\right|\right] + \mathbf{E}\left[\log|\mathbf{\Omega}|\right] - \mathbf{E}\left[\log\left|\mathbf{U}^H \hat{\mathbf{U}}\hat{\mathbf{U}}^H \mathbf{U}\right|\right] \\ &\leq \log\left|\mathbf{I}_{M_r} + \mathbf{E}\left[\mathbf{\Sigma}^{-2}\right]/\sigma_s^{\star 2}\right| + \log\left|\mathbf{I}_{M_r} + \sigma_a^{\star 2}\, \mathbf{E}\left[\mathbf{\Sigma}^2\right] \mathbf{E}\left[\mathbf{R}^H \mathbf{R}\right]\right| - \mathbf{E}\left[\log\left|\mathbf{Y}^H \mathbf{Y}\right|\right], \end{aligned} \quad (33)$$

where the second inequality follows from Jensen's inequality, the statistical independence between $\mathbf{\Sigma}$ and $\mathbf{U}^H \hat{\mathbf{U}}\hat{\mathbf{U}}^H \mathbf{U}$ [19], and Lemma 2.

For the first term in (33), since $\mathbf{E}\left[\mathrm{Tr}\left((\mathbf{H}_r \mathbf{H}_r^H)^{-1}\right)\right] = \mathbf{E}\left[\mathrm{Tr}(\mathbf{\Sigma}^{-2})\right] = M_r/(M_t - M_r)$ [19],

$$\begin{aligned} &\log\left|\mathbf{I}_{M_r} + \mathbf{E}\left[\mathbf{\Sigma}^{-2}\right] \frac{1}{\sigma_s^{\star 2}}\right| \\ \leq &M_r \log\left(\frac{\mathrm{Tr}\left(\mathbf{I}_{M_r} + \mathbf{E}\left[\mathbf{\Sigma}^{-2}\right]\right)}{M_r \sigma_s^{\star 2}} \right) \\ = &M_r \log\left(1 + \frac{1/\sigma_s^{\star 2}}{M_t - M_r}\right), \end{aligned}$$

where the first inequality is due to the fact that $|\mathbf{A}|^{\frac{1}{M_r}} \leq \mathrm{Tr}(\mathbf{A})/M_r$ for any $M_r \times M_r$ Hermitian positive semidefinite matrix $\mathbf{A}$. For the second term in (33), by Lemma 2 and the fact that $\mathbf{E}\left[\mathbf{\Sigma}^2\right] = M_t \mathbf{I}_{M_r}$ [19], one can show that

$$\begin{aligned} &\log\left|\mathbf{I}_{M_r} + \sigma_a^{\star 2}\, \mathbf{E}\left[\mathbf{\Sigma}^2\right] \mathbf{E}\left[\mathbf{R}^H \mathbf{R}\right]\right| \\ = &M_r \log\left(1 + \sigma_a^{\star 2} M_t D/M_r\right). \end{aligned}$$

The third term $-\mathbf{E}\left(\log\left|\mathbf{Y}^H \mathbf{Y}\right|\right)$ in (33) is, however, difficult to evaluate. By applying Jensen's inequality and Lemma 2, we have $-\mathbf{E}\left[\log\left|\mathbf{Y}^H \mathbf{Y}\right|\right] \geq -\log\left|\mathbf{E}\left[\mathbf{Y}^H \mathbf{Y}\right]\right| = -M_r \log(1 - D/M_r)$. We note that (by (30)) the right-hand side (RHS) of this inequality approaches zero for B sufficiently large; while the term $\log\left|\mathbf{Y}^H \mathbf{Y}\right| = \log\left|\mathbf{U}^H \hat{\mathbf{U}}\hat{\mathbf{U}}^H \mathbf{U}\right|$ is also close to zero with high probability for B sufficiently large due to random quantization codebook. We therefore make the approximation $-\mathbf{E}\left[\log\left|\mathbf{Y}^H \mathbf{Y}\right|\right] \approx -M_r \log(1 - D/M_r)$ for B sufficiently large. By substituting these results into (33), we obtain

$$\Delta R_M \lessapprox M_r \log\left(\frac{M_r + \sigma_a^{\star 2} M_t D}{M_r - D} \right) + M_r \log\left(1 + \frac{1/\sigma_s^{\star 2}}{M_t - M_r}\right). \quad (34)$$

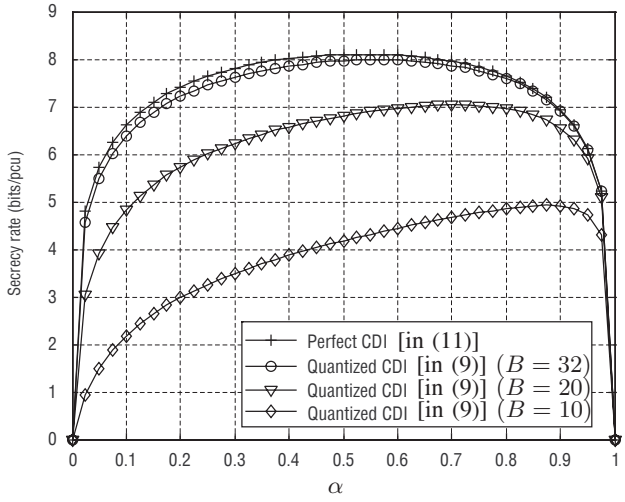

Fig. 2. Simulation results of the MISOSE secrecy rates [in (9) and (11)] versus the power allocation fraction α for $P = 25$ dB.

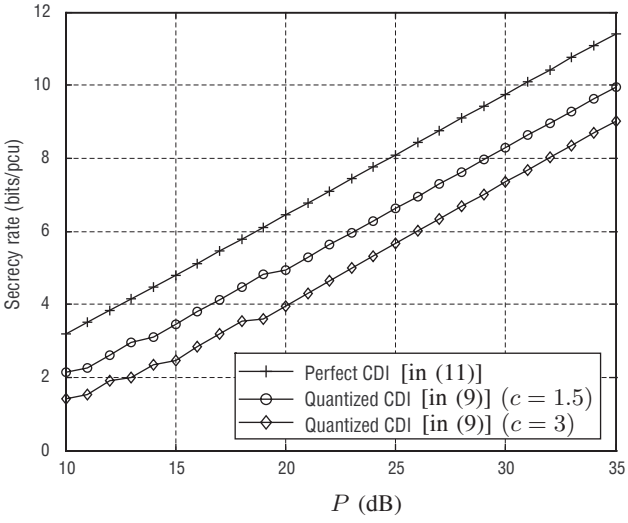

Fig. 3. Simulation results of the MISOSE secrecy rates [in (9) and (11)] versus the transmission power P for $\alpha = 0.5$. The number of feedback bits B is scaled with P according to the lower bound in (22).

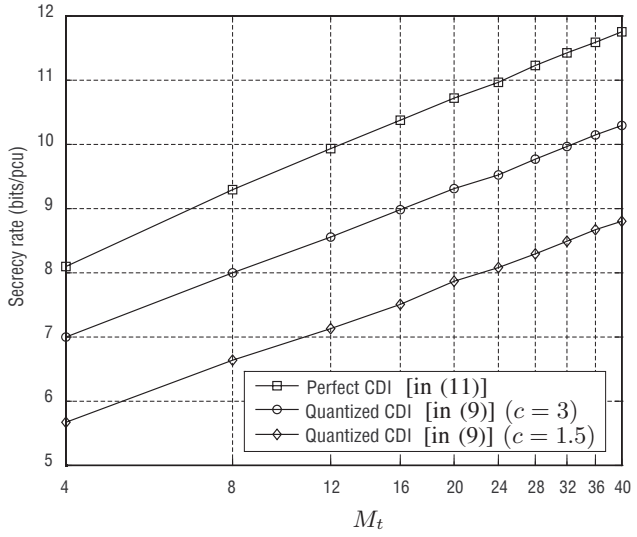

Fig. 4. Simulation results of the MISOSE secrecy rates [in (9) and (11)] versus the number of transmitter antenna M_t for $\alpha = 0.5$ and $P = 25$ dB. The number of feedback bits B is scaled with M_t according to the lower bound in (22).

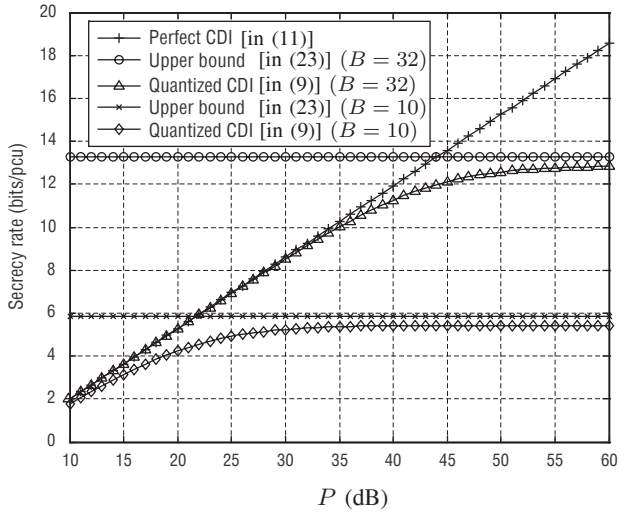

Fig. 5. Simulation results of the MISOSE secrecy rates [in (9) and (11)] versus the transmission power P for $\alpha = 0.9$. The MISOSE secrecy rate upper bound in (23) is also plotted.

By (34), to maintain a constant secrecy rate loss of c, i.e., $\Delta R_M \leq c$, it suffices to have

$$B \gtrapprox \widetilde{m} M_r \left[\log\left(1 + \frac{P(1-\alpha^\star_{\mathrm{p},M})(M_t/\widetilde{m})+1}{2^{c/M_r}(\alpha^\star_{\mathrm{p},M} P\widetilde{m})/(\alpha^\star_{\mathrm{p},M} P\widetilde{m}+1)-1}\right) + \log\left(\frac{\Gamma(\frac{1}{\widetilde{m}M_r})\Phi^{\frac{-1}{\widetilde{m}M_r}}}{\widetilde{m}M_r^2}\right)\right], \quad (35)$$

where $\widetilde{m} = M_t - M_r$. Note that the B in (35) scales as $\Omega(\log P)$ for fixed M_t and M_r.

VII. Simulation Results and Discussions

In this section, let us present simulation results to verify our analytical results. To be practical, we set the number of transmit antennas, i.e., M_t, to 4. Analytical results based on large M_t can be viewed as an approximation to the secrecy rate performance for practical values of M_t, whereas our simulation results match the analytical predictions well. Given a realization of $\mathbf{h}_r$, we used the numerical method in [9] to generate the associated quantized CDI $\hat{\mathbf{g}}_r$. This numerical method simulates the quantization procedure of random quantization codebook without generating a true codebook, thus saving a lot of computational time. Each simulation result was obtained by averaging over 10,000 channel realizations.

Figure 2 shows the simulation results of the MISOSE secrecy rate versus the power allocation fraction α for transmission power $P = 25$ dB. Both the secrecy rate with quantized CDI in (9) and that with perfect CDI in (11) are considered. As observed from this figure, with increased feedback bits B, the optimum α decreases from 0.9 (for $B = 10$) to 0.55 (for $B = 32$) and eventually gets close to 0.5 with perfect CDI. These results are consistent with Propositions 1 and 2. To

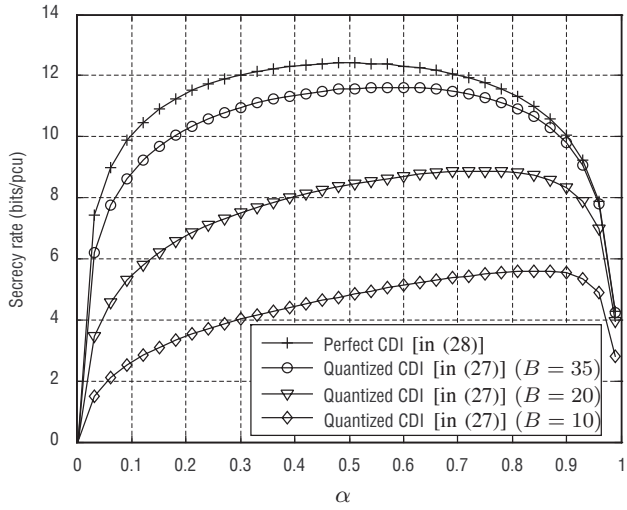

Fig. 6. Simulation results of MIMOME secrecy rates [in (27) and (28)] versus power allocation fraction α for $P = 25$ dB and $M_r = M_e = 2$.

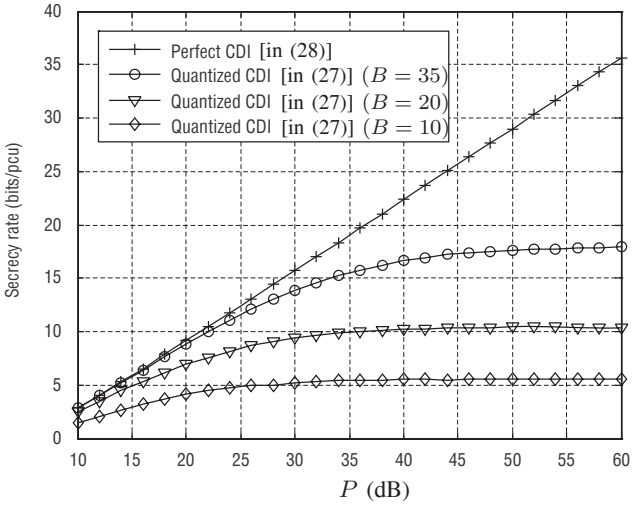

Fig. 7. Simulation results of the MIMOME secrecy rates [in (27) and (28)] versus the transmission power P for $\alpha = 0.9$ and $M_r = M_e = 2$.

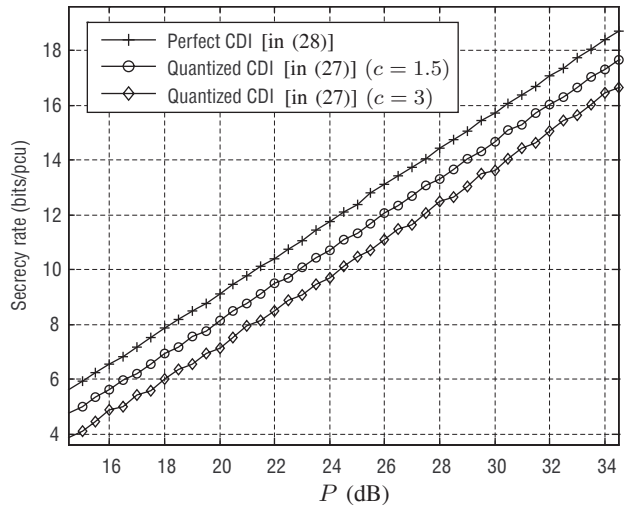

Fig. 8. Simulation results of the MIMOME secrecy rates [in (27) and (28)] versus the transmission power P for $\alpha = 0.5$ and $M_r = M_e = 2$. The number of feedback bits B is scaled with P according to the lower bound in (35).

verify the bit scaling law in Theorem 1, we show in Fig. 3 the simulation results of the MISOSE secrecy rate in (9) versus P for $\alpha = 0.5$. The number of feedback bits B is scaled according to the lower bound in (22) in order to maintain a constant secrecy rate loss c compared to the perfect CDI case. Note that as suggested by Fig. 2, the optimum $\alpha_{\rm p}^{\star}$ under perfect CDI in (22) is very close to 0.5. Here we examined the two cases of $c = 1.5$ and $c = 3$. One can see from this figure that, for $P = 30$ dB, the secrecy rate losses are respectively 1.463 bits/pcu and 2.408 bits/pcu, both of which are well controlled by the derived bit scaling law. In Fig. 4, we also verify the MISOSE secrecy rate in (9) versus M_t under $P = 25$ dB using (22). Again, the results agree with theoretical developments. In Fig. 5, we show the simulation results of the MISOSE secrecy rate in (9) versus P for $\alpha = 0.9$ and $B = 10$ and $B = 32$, respectively. The secrecy rate upper bound in (23) [see Theorem 2] is also simulated. As expected, the secrecy rate with quantized CDI is upper bounded, in sharp contrast to that with perfect CDI. Also, for large P, one can observe that the secrecy rate with quantized CDI in (9) approaches the upper bound in (23) (for both cases of $B = 10$ and $B = 32$), which coincides with our discussions in Section V and demonstrates that additional CQI at the transmitter only provides little gain on the achievable secrecy rate as P is large.

We show in Fig. 6 and Fig. 7 the MIMOME secrecy rate in (27) versus α and P, respectively. The numbers of antennas of the receiver and the eavesdropper are set as 2 ($M_r = M_e = 2$). Similar observations as in the MISOSE case can be seen from the two figures, that is, the optimum α decreases to around 0.5 when B increases, and the secrecy rate is upper bounded for finite B. The MIMOME bit scaling law in (35) is also examined, and the results are shown in Fig. 8. Although the lower bound in (35) is derived through some approximations, one can see from Fig. 8 that constant secrecy rate loss can still be maintained using this bit scaling law.

VIII. Conclusions

We have examined the impact of quantized CDI on the secrecy rate achievable with AN-assisted beamforming. In view of the AN leakage problem, we have analyzed the optimal power allocation strategy to maximize the achievable secrecy rate under quantized CDI. Moreover, to maintain a constant secrecy rate loss compared to the perfect CDI case, we have shown that the number of feedback bits B must be scaled logarithmically with the transmission power P for both the MISOSE and the MIMOME cases. The presented simulation results have confirmed our analytical results.

Acknowledgements

The authors would like to thank Mr. Meng-Hsi Chen and Mr. Chun-Yao Chen for the help on the MIMOME simulations.

APPENDIX

A. Proof of Lemma 1

Since $\mathbf{h}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_t})$, from [20, Chapter 5] we know that $|\mathbf{h}_e \hat{\mathbf{g}}_r|^2$ and $\|\mathbf{h}_e \mathbf{N}_{\hat{\mathbf{g}}_r}\|^2$ are independent and distributed as v_2 and v_{2M_t-2}, respectively, where v_k is chi-square distributed with degree of freedom k. Hence, it follows that

$$\begin{aligned}&\lim_{M_t\to\infty} \mathbf{E}\left[\log\left(1+\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}\frac{\alpha}{(1-\alpha)/(M_t-1)}\right)\right]\\ =&\lim_{M_t\to\infty} \mathbf{E}\left[\log\left(1+\frac{v_2}{v_{2M_t-2}/(M_t-1)}\frac{\alpha}{1-\alpha}\right)\right]\\ =&\mathbf{E}\left[\log\left(1+\frac{v_2}{2}\frac{\alpha}{1-\alpha}\right)\right] \end{aligned} \tag{36}$$

where the last equality follows by $v_{2M_t-2}/(2(M_t-1)) \xrightarrow{a.s.} 1$ as $M_t \to \infty$ from the law of large numbers (LLN). Then, equation (12) follows by evaluating the expectation over v_2. ■

B. Proof of Proposition 1

Let $d(\alpha^\star)$ and α_w be defined as in the Proposition statement, i.e., $d(\alpha^\star)$ is the RHS of (15) and α_w satisfies

$$\frac{1-\alpha_w}{\alpha_w} = d(\alpha_w).$$

Then, for $\alpha^\star \in (0,1)$ and ϵ such that $\frac{1-\alpha_w}{\alpha_w} > \epsilon > 0$, we have

Fact 1: $d(\alpha^\star) < 1$.

Fact 2: For $\alpha^\star < \alpha_w$, function

$$\frac{1-\alpha^\star}{\alpha^\star}\frac{1}{1+\epsilon} - d(\alpha^\star)$$

is strictly decreasing with respect to $\alpha^\star$.

Fact 3: For $\alpha^\star \geq \alpha_w + \epsilon_u$,

$$\frac{1-\alpha^\star}{\alpha^\star}(1+\epsilon) - d(\alpha^\star) \leq 0,$$

where $\frac{1-\alpha^\star}{\alpha^\star}(1+\epsilon) = d(\alpha^\star)|_{\alpha^\star=\alpha_w+\epsilon_u}$ since $\frac{1-\alpha^\star}{\alpha^\star}(1+\epsilon) - d(\alpha^\star)$ is strictly decreasing with respect to $\alpha^\star$.

Fact 1 can be easily shown by applying the following inequality [15]

$$\mathbf{E}_1(x)\exp(x) > \frac{1}{x+1},\ \forall x > 0. \tag{37}$$

in the definition of $d(\alpha^\star)$. For Fact 2, we want to show that

$$\begin{aligned}&\frac{1-\alpha^\star}{\alpha^\star}\frac{1}{1+\epsilon} - d(\alpha^\star)\\ =&\frac{1}{\alpha^*}\left(\frac{-\epsilon}{1+\epsilon} + \frac{1-\alpha^\star}{\alpha^\star}\mathbf{E}_1\left(\frac{1-\alpha^\star}{\alpha^\star}\right)\exp\left(\frac{1-\alpha^\star}{\alpha^\star}\right)\right) - \frac{1}{1+\epsilon}\end{aligned} \tag{38}$$

is strictly decreasing with respect to $\alpha^\star$ for $\alpha^\star < \alpha_w$. Note that when $x > 0$, $x\mathbf{E}_1(x)\exp(x)$ is strictly increasing with respect to x, since from (37) its derivative $-1+\mathbf{E}_1(x)\exp(x)(1+x) > 0$. Therefore, the function

$$\frac{-\epsilon}{1+\epsilon} + \frac{1-\alpha^\star}{\alpha^\star}\mathbf{E}_1\left(\frac{1-\alpha^\star}{\alpha^\star}\right)\exp\left(\frac{1-\alpha^\star}{\alpha^\star}\right)$$

is strictly decreasing with respect to $\alpha^\star$, for $\alpha^\star \in (0,1)$, and is larger than zero if $\alpha^\star < 1 - \frac{\epsilon}{1+\epsilon}$ which follows from (37). Moreover, since $\alpha^\star \in (0,1)$, the term $\frac{1}{\alpha^\star}$ is also greater than 0 and is strictly decreasing with respect to $\alpha^\star$. Therefore, for $\frac{1-\alpha_w}{\alpha_w} > \epsilon$ (or, equivalently, $1 - \frac{\epsilon}{1+\epsilon} > \alpha_w$), it follows that (38) strictly decreases with respect to $\alpha^\star$ for $\alpha^* < \alpha_w$. Fact 3 can be proven following similar procedures. Note that, since

$$\frac{1-\alpha^\star}{\alpha^\star}(1+\epsilon) - d(\alpha^\star)|_{\alpha^\star=\alpha_w} > \frac{1-\alpha^\star}{\alpha^\star} - d(\alpha^\star)|_{\alpha^\star=\alpha_w} = 0,$$

it follows that ϵ_u must be greater than 0 since (38) is strictly decreasing.

Given the facts above, let us first prove that $\alpha^\star > \alpha_w - \epsilon_l$. Define the event

$$A_\epsilon \triangleq \{\gamma_{\hat{\mathbf{g}}_r^\perp} P < \epsilon\}$$

for any $\epsilon > 0$. From (15) and the definition of $d(\alpha^\star)$ in the statement of Proposition, we have (39) in the top of the next page. As for equation (39), (a) follows from the fact that $0 \leq \gamma_{\hat{\mathbf{g}}_r^\perp} P < \epsilon$ under event A_ϵ and that $1-\alpha^\star < 1$, (b) follows by the definition of $\gamma_{\hat{\mathbf{g}}_r}$ in (14), and (c) is due to the fact that $\frac{B}{M_t-1} > \log(P/\epsilon)$ and that

$$|\mathbf{g}_r\hat{\mathbf{g}}_r^H|^2 = \cos^2\theta \geq 1-\delta = 1-2^{-\frac{B}{M_t-1}}$$

which follows from the QCA model in (4). Moreover, with the fact that $\|\mathbf{h}_r\|^2/M_t \xrightarrow{a.s.} 1$ as $M_t \to \infty$ (which follows by the LLN), we also have from (14) that

$$\gamma_{\hat{\mathbf{g}}_r^\perp} = \frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta < 2^{-\frac{B}{M_t-1}}$$

and, thus, $\Pr(A_\epsilon) \to 1$ as $M_t \to \infty$, for $B > (M_t-1)\log(P/\epsilon)$ (or equivalently $P2^{-\frac{B}{M_t-1}} < \epsilon$). Basically, the RHS in (39) will then approach $\frac{1-\alpha^\star}{\alpha^\star}\left(\frac{1}{1+\epsilon}\right)$ for M_t sufficiently large since $\|\mathbf{h}_r\|^2/M_t$ approaches 1 from the LLN and $1/M_t$ approaches 0. Then (39) becomes

$$d(\alpha^\star) > \frac{1-\alpha^\star}{\alpha^\star}\left(\frac{1}{1+\epsilon}\right). \tag{40}$$

Note that there must exist $\alpha^\star$ sufficiently small such that $\frac{1-\alpha^\star}{\alpha^\star}\frac{1}{1+\epsilon} - d(\alpha^\star) > 0$ (which follows by Fact 1) and that

$$\frac{1-\alpha^\star}{\alpha^\star}\frac{1}{1+\epsilon} - d(\alpha^\star)|_{\alpha^\star=\alpha_w} < \frac{1-\alpha^\star}{\alpha^\star} - d(\alpha^\star)|_{\alpha^\star=\alpha_w} = 0.$$

Then, by the strictly decreasing property given in Fact 2, there must exist $\epsilon_l > 0$ such that

$$\frac{1-\alpha^\star}{\alpha^\star}\frac{1}{1+\epsilon} = d(\alpha^\star)$$

for $\alpha^\star = \alpha_w - \epsilon_l$ and, thus, (40) holds for $\alpha^\star > \alpha_w - \epsilon_l$.

Now, let us prove that $\alpha^\star < \alpha_w + \epsilon_u$. Similarly, by (15), we can write

$$
\begin{aligned}
d(\alpha^\star) &> \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r} P(\gamma_{\hat{\mathbf{g}}_r^\perp} P+1)(1-\alpha^\star)}{(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+1)(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+\gamma_{\hat{\mathbf{g}}_r} P\alpha^\star+1)}\mathbf{1}_{\{A_\epsilon\}}\right] \\
&\overset{(a)}{>} \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r} P(1-\alpha^\star)}{(1+\epsilon)(\gamma_{\hat{\mathbf{g}}_r} P\alpha^\star+\epsilon+1)}\mathbf{1}_{\{A_\epsilon\}}\right] \overset{(b)}{=} \mathbf{E}\left[\frac{(1-\alpha^\star)}{\alpha^\star+\frac{\epsilon+1}{\|\mathbf{h}_r\|^2\cos^2\theta\cdot P}}\left(\frac{1}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right] \\
&\overset{(c)}{>} \mathbf{E}\left[\frac{(1-\alpha^\star)}{\alpha^\star+\frac{(\epsilon+1)/M_t}{\frac{\|\mathbf{h}_r\|^2}{M_t}(P-\epsilon)}}\left(\frac{1}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right]
\end{aligned} \tag{39}
$$

$$
\begin{aligned}
&d(\alpha^\star) = \\
&\mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r} P(\gamma_{\hat{\mathbf{g}}_r^\perp} P+1)(1-\alpha^\star)}{(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+1)(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+\gamma_{\hat{\mathbf{g}}_r} P\alpha^\star+1)}\mathbf{1}_{\{A_\epsilon\}}\right] \\
&+\mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r} P(\gamma_{\hat{\mathbf{g}}_r^\perp} P+1)(1-\alpha^\star)}{(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+1)(\gamma_{\hat{\mathbf{g}}_r^\perp} P(1-\alpha^\star)+\gamma_{\hat{\mathbf{g}}_r} P\alpha^\star+1)}\mathbf{1}_{\{A_\epsilon^c\}}\right].
\end{aligned} \tag{41}
$$

Note that the second term above can be written as (42) in the top of the next page, where the last inequality comes from the fact that $\alpha^\star > \alpha_w - \epsilon_l$ proven previously. Since $\Pr(A_\epsilon) \to 1$ as $M_t \to \infty$, the second term in (41) can be made arbitrarily small for M_t sufficiently large. Hence, it follows that, for M_t sufficiently large,

$$
\begin{aligned}
d(\alpha) &< \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r} P(1-\alpha^\star)(1+\epsilon)}{\gamma_{\hat{\mathbf{g}}_r} P\alpha^\star+1}\mathbf{1}_{\{A_\epsilon\}}\right] \\
&= (1+\epsilon)\ \mathbf{E}\left[\frac{(1-\alpha^\star)}{\alpha^\star+\frac{1}{\|\mathbf{h}_r\|^2\cos^2\theta\cdot P}}\mathbf{1}_{\{A_\epsilon\}}\right].
\end{aligned} \tag{43}
$$

Moreover, by following similar arguments in obtaining (40), (43) becomes

$$
d(\alpha^\star) < \frac{1-\alpha^\star}{\alpha^\star}(1+\epsilon)
$$

for M_t sufficiently large. It then follows, from Fact 3, that $\alpha^\star < \alpha_w + \epsilon_u$. We also plot functions $d(\alpha)$ and $(1-\alpha)/\alpha$ in Fig. 9. ■

C. Proof of Proposition 2

To prove this proposition, it suffices to show that, for any $\epsilon > 0$, if

$$
\alpha^\star < 1 - \frac{1}{\epsilon(\delta/2)P},
$$

then $\alpha^\star > 1/(1+\epsilon)$ for M_t sufficiently large. By the QCA model in (4) (where $\cos^2\theta \geq 1-\delta$) and by the definitions of $\gamma_{\hat{\mathbf{g}}_r}$ and $\gamma_{\hat{\mathbf{g}}_r^\perp}$ in (14), we have that

$$
\gamma_{\hat{\mathbf{g}}_r}/\gamma_{\hat{\mathbf{g}}_r^\perp} = (\cot^2\theta)(M_t-1) \geq (2^{\frac{B}{M_t-1}}-1)(M_t-1). \tag{44}
$$

Consider the event

$$
A_\epsilon \triangleq \{1/(\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P) < \epsilon\}.
$$

One can rearrange (15) to obtain (45) in the top of the next page, where the last inequality follows from (44) and the fact that $\alpha^\star \in (0,1)$. For $\alpha^\star < 1 - \frac{1}{\epsilon(\delta/2)P}$ (or, equivalently, when $\epsilon > \frac{1}{(\delta/2)(1-\alpha^\star)P}$), the event A_ϵ holds true whenever $\gamma_{\hat{\mathbf{g}}_r^\perp} > \delta/2$. Therefore we have

$$
\Pr(\gamma_{\hat{\mathbf{g}}_r^\perp} > \delta/2) \leq \Pr(A_\epsilon).
$$

By the LLN and (5), one can show that

$$
\Pr(\gamma_{\hat{\mathbf{g}}_r^\perp} > \delta/2) \to \Pr(\delta/2 \leq \sin^2\theta) = 1 - (1/2)^{M_t-1}.
$$

Then (45) becomes

$$
\begin{aligned}
&d(\alpha) > \\
&\frac{1}{(1+\epsilon)}\left(\alpha^\star + \frac{1+\epsilon}{(2^{\frac{B}{M_t-1}}-1)(M_t-1)}\right)^{-1}\left(1-\frac{1}{2^{M_t-1}}\right).
\end{aligned} \tag{46}
$$

Note that the RHS of (46) will approach $\frac{1}{\alpha^\star(1+\epsilon)}$ with $\eta \triangleq B/M_t > 0$ fixed and M_t sufficiently large. In order to have $d(\alpha^\star) > \frac{1}{\alpha^\star(1+\epsilon)}$, the value of $\alpha^\star$ must at least be larger than $1/(1+\epsilon)$ since $d(\alpha^\star) < 1$ by Fact 1 of Appendix B. We also plot functions $d(\alpha)$ and $1/\alpha$ in Fig. 9. ■

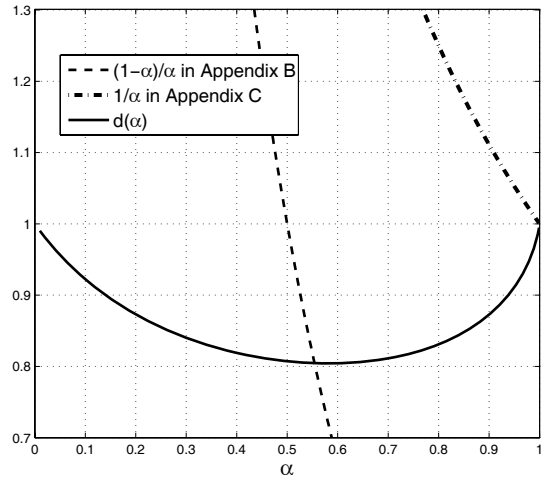

Fig. 9. Plots of functions related to the proofs of Proposition 1 and 2.

D. Proof of Theorem 1

It can be inferred that the ΔR in (20) can be further bounded as follows

$$
\Delta R \leq \mathbf{E}\left[\log\frac{\|\mathbf{h}_r\|^2\sin^2\theta\left(\frac{1-\alpha_{\mathrm{p}}^\star}{M_t-1}\right)P+1}{\cos^2\theta}\right]
$$

$$\mathbf{E}\Big[\frac{\gamma_{\hat{\mathbf{g}}_r^\perp}P(1-\alpha^\star)+(1-\alpha^\star)}{\gamma_{\hat{\mathbf{g}}_r^\perp}P(1-\alpha^\star)+1}\frac{\gamma_{\hat{\mathbf{g}}_r}P}{\gamma_{\hat{\mathbf{g}}_r}P\alpha^\star+(\gamma_{\hat{\mathbf{g}}_r^\perp}P(1-\alpha^\star)+1)}\mathbf{1}_{\{A_\epsilon^c\}}\Big] < \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r}P}{\gamma_{\hat{\mathbf{g}}_r}P\alpha^\star}\mathbf{1}_{\{A_\epsilon^c\}}\right] = \frac{1-\Pr(A_\epsilon)}{\alpha^\star} < \frac{1-\Pr(A_\epsilon)}{\alpha_w-\epsilon_l} \tag{42}$$

$$\begin{aligned} d(\alpha^\star) &> \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r}/\gamma_{\hat{\mathbf{g}}_r^\perp}}{(\gamma_{\hat{\mathbf{g}}_r}/\gamma_{\hat{\mathbf{g}}_r^\perp})\alpha^\star+\left(1+\frac{1}{\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P}\right)(1-\alpha^\star)}\left(\frac{\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P+1-\alpha^\star}{\gamma_{\hat{\mathbf{g}}_r^\perp}(1-\alpha^\star)P+1}\right)\mathbf{1}_{\{A_\epsilon\}}\right] \\ &> \mathbf{E}\left[\frac{\gamma_{\hat{\mathbf{g}}_r}/\gamma_{\hat{\mathbf{g}}_r^\perp}}{(\gamma_{\hat{\mathbf{g}}_r}/\gamma_{\hat{\mathbf{g}}_r^\perp})\alpha^\star+(1+\epsilon)(1-\alpha^\star)}\left(\frac{1+\epsilon(1-\alpha^\star)}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right] \geq \left(\alpha^\star+\frac{1+\epsilon}{(2^{\frac{B}{M_t-1}}-1)(M_t-1)}\right)^{-1}\frac{1}{(1+\epsilon)}\Pr(A_\epsilon), \end{aligned} \tag{45}$$

$$\begin{aligned} &+\mathbf{E}\left[\log\frac{1+\|\mathbf{h}_r\|^2\alpha_{\mathrm{p}}^\star P}{\|\mathbf{h}_r\|^2\alpha_{\mathrm{p}}^\star P}\right] \\ &\overset{(a)}{\leq}\log\mathbf{E}\left[\frac{\frac{\|\mathbf{h}_r\|^2(1-\alpha_{\mathrm{p}}^\star)}{M_t-1}P+1}{\cos^2\theta}-\frac{\|\mathbf{h}_r\|^2(1-\alpha_{\mathrm{p}}^\star)}{M_t-1}P\right] \\ &+\log\left(1+\frac{\mathbf{E}\left[1/\|\mathbf{h}_r\|^2\right]}{\alpha_{\mathrm{p}}^\star P}\right) \\ &\overset{(b)}{=}\log\left(\left(\frac{M_t(1-\alpha_{\mathrm{p}}^\star)}{M_t-1}P+1\right)\mathbf{E}\left[\sec^2\theta\right]-\frac{M_t(1-\alpha_{\mathrm{p}}^\star)}{M_t-1}P\right) \\ &+\log\left(1+\frac{1}{(M_t-1)\alpha_{\mathrm{p}}^\star P}\right) \end{aligned} \tag{47}$$

where (a) follows from Jensen's inequality [16], and (b) follows from the facts that $\mathbf{E}[\|\mathbf{h}_r\|^2]=M_t$, $\mathbf{E}\left(\frac{1}{\|\mathbf{h}_r\|^2}\right)=\frac{1}{M_t-1}$ [19], and that $\|\mathbf{h}_r\|$ and $\sin^2\theta$ are statistically independent [8]. By the QCA model where $\sin^2\theta<2^{-\frac{B}{M_t-1}}$, we have $\mathbf{E}\left[\sec^2\theta\right]\leq 1/\left(1-2^{\frac{-B}{M_t-1}}\right)$. Substituting this inequality into (47) gives rise to (1). Theorem 1 is then proved. ■

E. Proof of Theorem 2

Basically, we modify [2, Appendix B] to reach (50) and derive new results based on (50). First, let us define $\hat{y}_r^n=\{\hat{y}_r[i]\}_{i=1}^n$, $\hat{y}_e^n=\{\hat{y}_e[i]\}_{i=1}^n$ and $s^n=\{s[i]\}_{i=1}^n$. By Definition 1, for any $\varepsilon'>0$ and $n>n_0$, the achievable secrecy rate $\hat{R}$ can be upper bounded as

$$\begin{aligned} n\hat{R}-n\varepsilon' < H(w|\hat{y}_e^n,\mathbf{h}_r^n,\mathbf{h}_e^n) &\overset{(a)}{\leq} I(w;\hat{y}_r^n|\hat{y}_e^n,\mathbf{h}_r^n,\mathbf{h}_e^n)+n\varepsilon_1 \\ &\overset{(b)}{\leq} I(s^n;\hat{y}_r^n|\hat{y}_e^n,\mathbf{h}_r^n,\mathbf{h}_e^n)+n\varepsilon_1, \end{aligned} \tag{48}$$

where (a) follows from Fano's inequality for any $\varepsilon_1>0$ and sufficiently large n [2], and (b) follows from the Markov relation $w\leftrightarrow s^n\leftrightarrow(\hat{y}_r^n,\hat{y}_e^n)$. According to the channel model in (7) and (8), it can be shown [2] that

$$\begin{aligned} &I(s;\hat{y}_r|\hat{y}_e,\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)= \\ &[I(s;\hat{y}_r|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)-I(s;\hat{y}_e|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)]^+ \end{aligned} \tag{49}$$

for a certain fading state $(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)$. With (48) and (49), and by following [2] we have (50) in the top of the next page, where $\varepsilon_2=\varepsilon_1+\varepsilon'$ and $N(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)$ denotes the number of times for which the channel $(\mathbf{h}_r,\mathbf{h}_e)$ is in fading state $(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)$ over the n channel uses.

Notice that, for a given channel state $(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)$, (7) and (8) is equivalent to a scalar Gaussian wiretap channel, and therefore $[I(s;\hat{y}_r|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)-I(s;\hat{y}_e|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)]^+$ in (50) is maximum with Gaussian s (see [21]). Then, as [2], we denote the average power of $s[i]$ and $\mathbf{a}[i]$ over the channels $(\mathbf{h}_r[i],\mathbf{h}_e[i])=(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)$ in n channel uses as P_s^n and P_a^n, respectively. It is easy to see that

$$\begin{aligned} &[I(s;\hat{y}_r|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)-I(s;\hat{y}_e|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)]^+ \\ &\leq\left[\log\left(1+\frac{|\tilde{\mathbf{h}}_r\hat{\mathbf{g}}_r^H|^2P_s^n}{\|\tilde{\mathbf{h}}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2P_a^n+1}\right)-\log\left(1+\frac{|\tilde{\mathbf{h}}_e\hat{\mathbf{g}}_r^H|^2P_s^n}{\|\tilde{\mathbf{h}}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2P_a^n}\right)\right]^+ \\ &\leq\left(\log\frac{P_a^n+\frac{|\tilde{\mathbf{h}}_r\hat{\mathbf{g}}_r^H|^2}{\|\tilde{\mathbf{h}}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s^n}{P_a^n+\frac{|\tilde{\mathbf{h}}_e\hat{\mathbf{g}}_r^H|^2}{\|\tilde{\mathbf{h}}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s^n}\right)^+. \end{aligned}$$

By substituting this inequality into (50) and by following steps in [2], we have

$$\begin{aligned} \hat{R}-\varepsilon_2 &\leq\int\int\left(\log\frac{P_a^n+\frac{|\tilde{\mathbf{h}}_r\hat{\mathbf{g}}_r^H|^2}{\|\tilde{\mathbf{h}}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s^n}{P_a^n+\frac{|\tilde{\mathbf{h}}_e\hat{\mathbf{g}}_r^H|^2}{\|\tilde{\mathbf{h}}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s^n}\right)^+\frac{N(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)}{n}d\tilde{\mathbf{h}}_r d\tilde{\mathbf{h}}_e \\ &=\mathbf{E}\left[\left(\log\frac{P_a+\frac{|\mathbf{h}_r\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s}{P_a+\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}P_s}\right)^+\right], \end{aligned} \tag{51}$$

where P_s and P_a denote the limits of P_s^n and P_a^n for $n\to\infty$, respectively, and the last equality follows from the ergodicity of the channel [2]. Notice that, for $\frac{|\mathbf{h}_r\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}\geq\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}$,

$$\begin{aligned} &\log\left(\frac{P_a+P_s\,|\mathbf{h}_r\hat{\mathbf{g}}_r^H|^2/\|\mathbf{h}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}{P_a+P_s\,|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2/\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}\right) \\ &\leq\left[\log\frac{|\mathbf{h}_r\hat{\mathbf{g}}_r^H|^2\;\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}{\|\mathbf{h}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2\;|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}\right]^+, \end{aligned} \tag{52}$$

since $(a+b)/(a+d)\leq b/d$ for $a\geq 0$ and $b\geq d\geq 0$. The bound also holds true for $\frac{|\mathbf{h}_r\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}<\frac{|\mathbf{h}_e\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{h}_e\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}$ since the left-hand-side (LHS) of the above inequality is negative in this case. This leads to the result in (23). By following

$$\hat{R} \leq \int\int [I(s;\hat{y}_r|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e) - I(s;\hat{y}_e|\mathbf{h}_r=\tilde{\mathbf{h}}_r,\mathbf{h}_e=\tilde{\mathbf{h}}_e)]^+ \frac{N(\tilde{\mathbf{h}}_r,\tilde{\mathbf{h}}_e)}{n} d\tilde{\mathbf{h}}_r d\tilde{\mathbf{h}}_e + \varepsilon_2, \tag{50}$$

the arguments in [2], one can show that the bound also holds true for finite n. Note that the bound holds regardless of the power allocation under any channel state. Thus the bound is valid not only for the no CQI case with fixed power allocation for all channel states, but also valid even if the transmitter has perfect knowledge of the CQI and is able to perform power allocation across time-varying channel states. ■

F. Proof of Corollary 1

Following arguments as those in the proof of Lemma 1, the upper bound in (23) equals to

$$\begin{aligned}
&\mathbf{E}\left[\left(\log\frac{|\mathbf{g}_r\hat{\mathbf{g}}_r^H|^2}{\|\mathbf{g}_r\mathbf{N}_{\hat{\mathbf{g}}_r}\|^2}\frac{v_{2M_t-2}}{v_2}\right)^+\right]\\
=&\mathbf{E}\left[\left(\log\cot^2\theta\frac{v_{2M_t-2}}{v_2}\right)^+\right]\\
=&\mathbf{E}\left[\log\left(\cot^2\theta\frac{v_{2M_t-2}}{v_2}\right)\right]\\
&-\mathbf{E}\left[\log\left(\cot^2\theta\frac{v_{2M_t-2}}{v_2}\right)\mathbf{1}_{\{\cot^2\theta\ v_{2M_t-2}<v_2\}}\right].
\end{aligned} \tag{53}$$

With M_t sufficiently large and define $F(2,2M_t-2) \triangleq \frac{v_2/2}{v_{2M_t-2}/(2M_t-2)}$, by the QCA model in (4), we have (54) in the top of next page where $\varepsilon_1 > 0$ goes to zero as M_t goes to infinity. We now prove inequality (a) in (54). By the LLN, the distribution of $F(2,2M_t-2)$ will approach that of v_2 as M_t goes to infinity. Then the LHS of (a) in (54) equals to (55) in the top of the next page, where last inequality holds by choosing M_t sufficiently large since $\mathbf{E}_1(x)$ is a non-negative and strictly decreasing function with $\lim_{x\to\infty}\mathbf{E}_1(x)=0$.

For M_t and P sufficiently large and by applying the inequality $\mathbf{E}_1(x)e^x < \ln(\frac{1+x}{x})$ [15] along with (12), the achievable secrecy rate in (9) can be lower-bounded by

$$\begin{aligned}
\hat{R}(\alpha)>&\mathbf{E}\left[\log\left(1+\frac{\|\mathbf{h}_r\|^2\cos^2\theta\cdot\frac{\alpha}{1-\alpha}}{\frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta+\frac{1}{(1-\alpha)P}}\right)\right]\\
&-\log\left(\frac{1}{1-\alpha}\right)-\varepsilon_2,
\end{aligned} \tag{56}$$

where $|\varepsilon_2|$ can be arbitrarily small. Let us define the event

$$A_\epsilon \triangleq \left\{\frac{1}{\epsilon} < \frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta(1-\alpha)P\right\},$$

where $\epsilon > 0$. Then the first term of the RHS of (56) can be further lower-bounded as

$$\begin{aligned}
&\mathbf{E}\left[\log\left(1+\frac{\|\mathbf{h}_r\|^2\cos^2\theta\cdot\frac{\alpha}{1-\alpha}}{\frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta+\frac{1}{(1-\alpha)P}}\right)\right]\\
>&\mathbf{E}\left[\left(\log\left(1+\frac{\|\mathbf{h}_r\|^2\cos^2\theta\cdot\frac{\alpha}{1-\alpha}}{\frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta+\frac{1}{(1-\alpha)P}}\right)\right)\mathbf{1}_{\{A_\epsilon\}}\right]\\
>&\mathbf{E}\left[\log\left(\frac{(M_t-1)\cot^2\theta\cdot\frac{\alpha}{1-\alpha}}{1+\frac{1}{\frac{\|\mathbf{h}_r\|^2}{M_t-1}\sin^2\theta(1-\alpha)P}}\right)\mathbf{1}_{\{A_\epsilon\}}\right]\\
>&\mathbf{E}\left[\log\left(\frac{(M_t-1)\cot^2\theta\cdot\frac{\alpha}{1-\alpha}}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right]\\
=&\mathbf{E}\big[\log\big(\cot^2\theta\big)\big]-\mathbf{E}\big[\log\big(\cot^2\theta\big)\mathbf{1}_{\{A_\epsilon^c\}}\big]\\
&+\mathbf{E}\left[\log\left(\frac{(M_t-1)\frac{\alpha}{1-\alpha}}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right]\\
>&\mathbf{E}\left[\log\left((M_t-1)\cot^2\theta\cdot\frac{\alpha}{1-\alpha}\right)\right]-\varepsilon_2',
\end{aligned} \tag{57}$$

where $\varepsilon_2' > 0$ goes to zero when M_t goes to infinity. The last inequality of (57) come as follows. Apply the LLN on $\frac{|\mathbf{h}_r\|^2}{M_t-1}$, we know that event A_ϵ happens when $\{\cot^2\theta < \epsilon(1-\alpha)P-1\}$ for M_t large enough. If we choose P satisfying $\epsilon(1-\alpha)P > 2/\delta$, from the probability density function (PDF) of $\cot^2\theta$ [7]

$$\begin{aligned}
&\mathbf{E}\left[\log\big(\cot^2\theta\big)\mathbf{1}_{\{A_\epsilon^c\}}\right]\\
=&\int_{\epsilon(1-\alpha)P-1}^{\infty}\log x\frac{2^B(M_t-1)}{(x+1)^{M_t}}dx\\
<&\int_{\frac{2}{\delta}-1}^{\infty}\log x\frac{2^B(M_t-1)}{(x+1)^{M_t}}dx
\end{aligned} \tag{58}$$

where $\delta = 2^{-B/(M_t-1)}$. The RHS of the last inequality of (58) is upper-bounded by

$$\int_{\frac{2}{\delta}-1}^{\infty}\log(x+1)\frac{2^B(M_t-1)}{(x+1)^{M_t}}dx = \frac{1}{2^{M_t-1}}\left[\frac{B+\log e}{M_t-1}+1\right].$$

With large M_t and fixed $\eta \triangleq B/M_t$, the RHS of the above inequality can be made arbitrary small. Moreover, using the methods under (45), $\Pr(\mathrm{A}_\epsilon) \geq 1-(1/2)^{M_t-1}$, then

$$\begin{aligned}
&\mathbf{E}\left[\log\left(\frac{(M_t-1)\frac{\alpha}{1-\alpha}}{1+\epsilon}\right)\mathbf{1}_{\{A_\epsilon\}}\right]\\
\geq&\log\left(\frac{(M_t-1)\frac{\alpha}{1-\alpha}}{1+\epsilon}\right)-\frac{\log\left(\frac{(M_t-1)\frac{\alpha}{1-\alpha}}{1+\epsilon}\right)}{2^{M_t-1}},
\end{aligned} \tag{59}$$

With large M_t, the second term in the RHS of (59) approaches to zero. Then from (58) and (59), we can prove the final inequality of (57).

Finally, from (53), (54), (56), (57) and by applying the LLN on $\frac{v_{2M_t-2}}{2M_t-2}$, the difference between the upper bound in (23) and

$$
\begin{aligned}
&-\mathbf{E}\left[\log\left(\cot^2\theta\frac{v_{2M_t-2}}{v_2}\right)\mathbf{1}_{\{\cot^2\theta\, v_{2M_t-2}<v_2\}}\right] < \mathbf{E}\left[\log\left(\frac{v_2}{v_{2M_t-2}(2^{\frac{B}{M_t-1}}-1)}\right)\mathbf{1}_{\{(2^{\frac{B}{M_t-1}}-1)\, v_{2M_t-2}<v_2\}}\right]\\
&=\mathbf{E}\left[\log\left(\frac{F(2,M_t-2)}{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}\right)\mathbf{1}_{\{(M_t-1)(2^{\frac{B}{M_t-1}}-1)<F(2,M_t-2)\}}\right]\overset{(a)}{<}\varepsilon_1, \qquad (54)
\end{aligned}
$$

$$
\begin{aligned}
&\int_{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}^{\infty}\log\left(\frac{x}{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}\right)\frac{1}{2}\exp(-x/2)dx\\
&=-(\log e)\,\mathbf{E}_1\left(\frac{x}{2}\right)-\log\left(\frac{x}{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}\right)\exp(-x/2)\Big|_{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}^{\infty}\\
&=-\log e\left(\lim_{x\to\infty}\mathbf{E}_1\left(\frac{x}{2}\right)-\mathbf{E}_1\left(\frac{(M_t-1)(2^{\frac{B}{M_t-1}}-1)}{2}\right)\right)<\varepsilon_1, \qquad (55)
\end{aligned}
$$

the achievable secrecy rate in (9) is no greater than $|\log\alpha|-\mathbf{E}[\log(v_2/2)]+\varepsilon$ for M_t and P large enough, where $\varepsilon=|\varepsilon_1+\varepsilon_2+\varepsilon_2'|$. Since $\mathbf{E}[\log(v_2/2)]$ can be numerically found as -0.832, it then concludes the proof. ■

REFERENCES

[1] A. Wyner, "The wiretap channel," *Bell Syst. Technical J.*, vol. 54, pp. 1355-1387, 1975.

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

[3] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470-2492, 2008.

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

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

[6] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," submitted to *IEEE Trans. Inf. Theory*, Mar. 2009.

[7] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478-1491, 2007.

[8] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045-5060, 2006.

[9] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1473-1482, 2008.

[10] 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 with artificial noise," in *Proc. IEEE ISIT*, Seoul, Korea, June 2009, pp. 2351-2355.

[11] S.-C. Lin, T.-H. Chang, Y.-W. Hong, and C.-Y. Chi, "On the impact of quantized channel direction feedback in multiple-antenna wiretap channels," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010.

[12] S. Jafar and A. Goldsmith, "Isotropic fading vector broadcast channels: the scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 848-857, 2005.

[13] M. Yuksel and E. Erkip, "Diversity-multiplexing tradeoff for the multiple-antenna wire-tap channel," submitted to *IEEE Trans. Wireless Commun.*, June 2009.

[14] K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite-rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2562-2579, 2003.

[15] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1964.

[16] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.

[17] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2001.

[18] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108-1123, 2008.

[19] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. now Publisher Inc., 2004.

[20] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*. Marcel Dekker, 1992.

[21] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451-456, 1978.

Shih-Chun Lin (M'08) received the B.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2000 and 2007, respectively. He was a Visiting Student at The Ohio State University, Columbus, in 2007. After serving his military duty in 2008, Dr. Lin is a postdoctoral research associate at the Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan now. He also served as a TPC member of IEEE VTC 2010-Spring. His research interests include coding/information theory, communications, and signal processing.

Tsung-Hui Chang (S'07-M'08) received his B.S. degree in electrical engineering and his Ph.D. degree in communications engineering from the National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2003 and 2008, respectively. During September 2006 and February 2008, he was an exchange Ph.D. student of University of Minnesota, Minneapolis, Minnesota, USA. Currently, he is a postdoctoral research fellow with the Institute of Communications Engineering, NTHU. His research interests are widely in wireless communications, digital signal processing and convex optimization and its applications.

Ya-lan Liang received the B.S. degree in the Electrical Engineering from National Cheng Kung University in 2007 and M.S. degree in the Communication Engineering from Tsing Hua University in 2009. Currently, she is a research engineer at Hon Hai Precision Ind. Co., Ltd. Her research interests include MIMO system and Wimax communication.

Y.-W. Peter Hong (S'01-M'05) received his B.S. degree in Electrical Engineering from National Taiwan University, Taipei, Taiwan, in 1999, and his Ph.D. degree in Electrical Engineering from Cornell University, Ithaca, NY, in 2005. He joined the Institute of Communications Engineering and the Department of Electrical Engineering at National Tsing Hua University, Hsinchu, Taiwan, in Fall 2005, where he is now an Associate Professor. He was also a visiting scholar at the University of Southern California during June-August of 2008. His research interests include cooperative communications, distributed signal processing for sensor networks, physical layer secrecy, and PHY-MAC cross-layer designs for wireless networks. Dr. Hong received the best paper award for young authors from the IEEE IT/COM Society Taipei/Tainan chapter in 2005, the best paper award among unclassified papers in MILCOM 2005, the Junior Faculty Research Award from the College of EECS and from National Tsing Hua University in 2009 and 2010, respectively, and the Outstanding Teaching Award from the College of EECS in 2010. In 2010, he also received the Asia-Pacific Outstanding Young Researcher Award from the IEEE Communication Society. He is a co-editor (along with A. Swami, Q. Zhao, and L. Tong) of the book entitled *Wireless Sensor Networks: Signal Processing and Communications Perspectives* (John-Wiley & Sons, 2007), and is a coauthor (along with W.-J. Huang and C.-C. Jay Kuo) of the book entitled *Cooperative Communications and Networking: Technologies and System Design*. Dr. Hong has served as Publication Co-Chair and TPC Track Co-Chair of VTC2010-Spring for the track on "Cognitive Radio and Cooperative Communications" and also as Publicity Co-Chair of ISITA/ISSSTA 2010. Dr. Hong is also a guest editor of EURASIP Special Issue on Cooperative MIMO Multicell Networks and of IJSNET Special Issue on Advances in Theory and Applications of Wireless, Ad Hoc, and Sensor Networks. He is also currently serving as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Chong-Yung Chi (S'83-M'83-SM'89) received the Ph.D. degree in Electrical Engineering from the University of Southern California, Los Angeles, California, in 1983. From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, California. He has been a Professor with the Department of Electrical Engineering since 1989 and the Institute of Communications Engineering (ICE) since 1999 (also the Chairman of ICE during 2002-2005), National Tsing Hua University, Hsinchu, Taiwan. He has published more than 160 technical papers, including more than 50 journal papers (mostly in IEEE Trans. Signal Processing), two book chapters, and more than 100 peer-reviewed conference papers, as well as a graduate-level textbook, *Blind Equalization and System Identification* (Springer-Verlag, 2006). His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, biomedical and hyperspectral image analysis.

Dr. Chi is a senior member of IEEE. He has been a Technical Program Committee member for many IEEE sponsored and co-sponsored workshops, symposiums and conferences on signal processing and wireless communications, including Co-organizer and General Co-chairman of 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), and Co-Chair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008 & Lead Co-Chair of SPC Symposium, ChinaCOM 2009. He is currently serving as Track Chair for MIMO, Signal Processing, and Smart in Antennas, 2011 IEEE Radio and Wireless Symposium in Radio and Wireless Week (RWW) 2011. He was an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING (5/2001-4/2006), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (1/2006-12/2007), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (1/2008-12/2009), Associate Editor of IEEE SIGNAL PROCESSING LETTERS (6/2006-5/2010), and a member of Editorial Board of *EURASIP Signal Processing Journal* (6/2005-5/2008), and an editor (7/2003-12/2005) as well as a Guest Editor (2006) of *EURASIP Journal on Applied Signal Processing*. He was a member of IEEE Signal Processing Committee on Signal Processing Theory and Methods (2005-2010). Currently he is a member of IEEE Signal Processing Committee on Signal Processing for Wireless Communications and Networking.