Artificial-Noise-Aided Secure Transmission Scheme with Limited Training and Feedback Overhead

Jianwei Hu, *Student Member, IEEE,* Yueming Cai, *Senior Member, IEEE,* Nan Yang, *Member, IEEE,* Xiangyun Zhou, *Member, IEEE,* and Weiwei Yang, *Member, IEEE*

***Abstract*—We design a novel artificial-noise-aided secure on-off transmission scheme in a wiretap channel. We consider a practical scenario where the multi-antenna transmitter only obtains partial channel knowledge from the single-antenna receiver through limited training and feedback but has no channel knowledge about the single-antenna eavesdropper. In the design, we first propose a three-period block transmission protocol to capture the practical training and quantization features. We then characterize the statistics of the received signal-to-noise ratios (SNRs) at the receiver and the eavesdropper. Under the secrecy outage constraint, we exploit the on-off scheme to perform secure transmission and derive a closed-form expression for the secrecy throughput. Moreover, we investigate the optimization problem of maximizing the secrecy throughput by proposing an iterative algorithm to determine the optimal power allocation between the information signal and artificial noise, as well as the optimal codeword transmission rate. Furthermore, we define the net secrecy throughput (NST) which takes the signaling overhead into account and address the problem of optimally allocating the block resource to the training and feedback overhead. Numerical results clearly demonstrate how the optimal signaling overhead changes with the number of transmit antennas, and there exists an optimal number of antennas that maximizes the NST.**

***Index Terms*—Secure transmission design, artificial noise, training, feedback, net secrecy throughput.**

I. Introduction

THE SECURITY issue of data transmission is an increasingly pressing concern in the design and development of wireless applications, e.g., financial information, electronic media, medical records, and customer files in the fifth generation (5G) network [1]. Indeed, the future wireless network connecting millions of wireless devices would face serious and persistent security threats if reliable security mechanism is not established [2].

Traditional security is achieved based on the assumption that the computational resources of the adversaries are insufficient to break specific public-key cryptography in polynomial time [3]. However, it would be completely compromised when the quantum computers of non-trivial size become reality. Against this backdrop, the information-theoretic security approach has been proposed and developed as a promising complement to the cryptographic techniques. The seminal work in this area can be traced back to [4] in 1975, which concluded that perfect secrecy can be achieved at the physical layer by taking channel impairments into account even when the adversaries know the details of the codes employed. This conclusion has laid the foundation of the so-called *physical layer security* and triggered numerous research studies to investigate different physical layer methods for security enhancement. Some exemplary methods are transmit beamforming [5–7], antenna selection [8–10], cooperation techniques [11–13], and artificial-noise-aided transmission [14–17].

A. Related Work

A common assumption adopted by the previous studies on physical layer security is that the transmitter can obtain perfect channel state information (CSI) of the legitimate receiver and/or the eavesdropper. Such idealized conditions provide us attractive advantages of the physical layer security. However, due to the practical issues of training and channel feedback, it is controversial to assume that the transmitter perfectly knows the legitimate receiver's channel, let alone the eavesdropper's channel [18]. For example, in practical frequency division duplex (FDD) systems, prior to data communication the receiver first needs to estimate its channel by detecting the training symbols emitted from the transmitter. Then a feedback link is utilized for the receiver to convey the estimated CSI back to the transmitter, which forms the closed-loop transmission. Since the rate of the feedback link is usually limited, the receiver has to quantize the estimated CSI before feeding it back. This means that the accuracy of channel estimation and quantization at the receiver affects the quality of the CSI at the transmitter (CSIT). Due to this, considerable research efforts have been devoted to designing practical secure transmission schemes by considering imperfect CSIT caused by the practical training and feedback issues.

Considering the imperfect channel estimation at receivers, the optimal tradeoff between the energy used for training and data signals was characterized for the multiple-input single-output single-eavesdropper (MISOSE) wiretap channel in [19]. Moreover, the on-off mechanism was adopted to design an practical transmission scheme in the presence of channel

This work of J. Hu, Y. Cai and W. Yang was supported by the National Natural Science Foundation of China (No. 61371122). The work of N. Yang and X. Zhou was supported by the Australian Research Council Discovery Project (DP150103905).

J. Hu, Y. Cai and W. Yang are with the College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China (e-mail: hujianwei1990@yeah.net, caiym@vip.sina.com, wwyang1981@163.com).

N. Yang and X. Zhou are with the Research School of Engineering, Australian National University, Canberra, ACT 0200, Australia (email: {nan.yang, xiangyun.zhou}@anu.edu.au).

Digital Object Identifier 10.1109/TWC.2016.2621040

estimation errors [20]. On the other hand, considering the scenario with limited feedback, [21] studied the secrecy outage probability when the predefined codebook is known to both transmitter and receiver. Focusing on the artificial-noise-aided beamforming scheme, [22] and [23] optimized the transmission to guarantee secrecy with quantized channel feedback for fast and slow fading channels, respectively. For block fading channels, we note that dedicating a significant amount of training and feedback overhead offers the high quality of CSIT, but reduces the effective time resource used for data transmission [24, 25]. Motivated by this, [26] studied the non-trivial tradeoff in wiretap channels, in which the effective ergodic secrecy rate (ESR) was defined as the performance metric and the maximization of the effective ESR was investigated. We clarify that the ESR may not be an appropriate performance metric in the system with stringent delay constraints. Actually, outage-based characterizations, which measure systems with probabilistic formulations, become more appropriate [27]. This motives us to address the following problem in this paper: "*How to achieve the maximum net secrecy throughput (NST) under the secrecy outage constraint in quasi-static block fading channels?*" Here, the NST is defined as the secrecy throughput excluding the training and feedback overhead.

B. Our Contributions

In this paper, we focus on the MISOSE wiretap channel over quasi-static block fading, where the communication between a multi-antenna transmitter and a single-antenna legitimate receiver is overheard by a single-antenna malicious eavesdropper. Importantly, we adopt a *practical* CSI assumption that the transmitter only obtains the partial knowledge of the transmitter-receiver channel through a finite amount of training and feedback overhead. In contrast to the analog feedback used in [26], in this work we adopt the digital (or equivalently, quantized) feedback, which is due to the fact that the digital feedback has a low complexity and offers a substantial advantage over the analog feedback [24, 25].

In our considered wiretap channel where the instantaneous knowledge of the transmitter-eavesdropper channel is not available at the transmitter, the secrecy outage events will inevitably happen. To guarantee a target security level, we propose an artificial-noise-aided secure on-off transmission scheme under the secrecy outage constraint[1]. We then derive the closed-form expression for the secrecy throughput and investigate the optimization problem of the secrecy throughput maximization. Moreover, by adopting the NST as our performance metric, we examine the optimum fraction of resource that needs to be dedicated to the training and feedback overhead. The key contributions of this paper are summarized as follows:

1) We propose a practical three-period block transmission protocol to help the transmitter obtain the legitimate receiver's CSI and perform secure transmission. The proposed protocol divides each block into three periods, namely, the training period, the digital feedback period, and the data transmission period. We clarify that this protocol stands as a perfectly suitable solution to the realistic system, e.g., the FDD system, since it captures the essential system features.
2) We characterize new statistics of the received signal-to-noise ratio (SNR) at the legitimate receiver under limited training and feedback overhead. Such statistics enables us to precisely characterize the secrecy throughput performance. We state that the newly derived statistics has never been presented in the literature, and serves as a useful formula to examine the system performance in similar applications, e.g., multi-user multi-antenna communication with finite signaling overhead.
3) We develop new guidelines to determine the wiretap code parameters for secure transmission, where the rate redundancy is designed to keep the required security level, and the codeword transmission rate is designed as an constant parameter to optimize. Using this fixed-rate encoding strategy, we derive the closed-form expression for the secrecy throughput. Then we investigate the optimization problem of secrecy throughput maximization and propose an iterative algorithm to find the optimal power allocation ratio and codeword transmission rate.
4) We adopt the NST as a more appropriate secrecy performance metric and examine the optimum fraction of block resource allocated to the training symbols and feedback bits. We numerically find that the optimal fraction of the signaling overhead is approximately proportional to the number of the transmit antennas. We further find that increasing the number of actively used transmit antennas does not always improve the NST. Instead, there exists an optimal number of transmit antennas that maximizes the NST. As such, when we have enough transmit antennas, we need to carefully determine a suitable number of antennas according to the system parameters.

C. Organization

The remainder of this paper is organized as follows. In Section II, we provide an overview of the considered MISOSE scenario. In Section III, we characterize the statistics of the received SNRs. In Section IV, we design the transmission scheme under secrecy outage constraint. In Section V, we use the NST as a more practical metric to examine the optimum fractional overhead. Finally, we present our numerical simulations and main findings in Section VI and VII, respectively.

Notation: Matrices and column vectors are denoted by uppercase and lowercase boldface letters. A circularly symmetric complex Gaussian random variable z with variance σ^2 is denoted as $z \sim \mathcal{CN}(0, \sigma^2)$. A Gamma-distributed random variable x with parameters (a, b) is denoted as $x \sim Gamma(a, b)$. $\mathbf{I}_N$ stands for an identity matrix of size N. $(\cdot)^\dagger$ stands for conjugate transpose operation. $null(\cdot)$ stands for spanning the null space of matrix. $|\cdot|$ and $\|\cdot\|$ represent the norm of scalar

[1]The on-off transmission scheme enables the transmitter to decide whether or not to transmit according to the knowledge about the CSI of the transmitter-receiver channel and/or the transmitter-eavesdropper channel. With this decision, the reliability and security performance can be enhanced [27, 28].

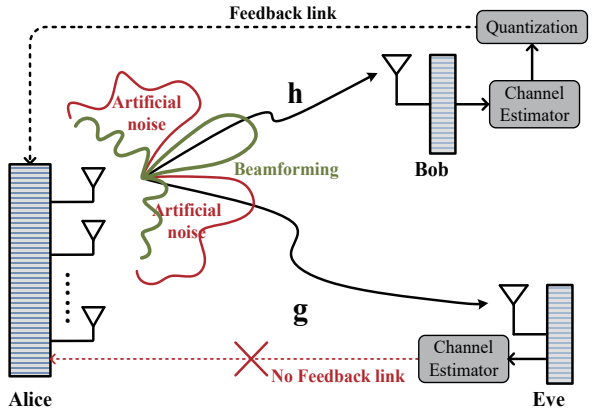

Fig. 1. A MISOSE wiretap channel with channel estimation error and limited digital feedback.

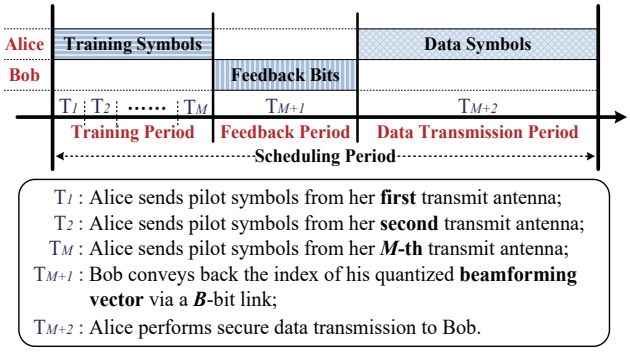

Fig. 2. Scheduling period partition of a specific coherence block in the proposed three-period block transmission protocol.

and vector, respectively. $\log_2(\cdot)$ and $\ln(\cdot)$ represent the base 2 logarithm and natural logarithm, respectively.

II. System Model

We consider a MISOSE wiretap channel, as illustrated in Fig. 1, where the transmitter, Alice, communicates with a desired receiver, Bob, in the presence of a passive eavesdropper, Eve. We assume that Alice is equipped with M antennas, while Bob and Eve are equipped with a single antenna each. Throughout this paper, we refer to the Alice-Bob link as the main channel and the Alice-Eve link as the eavesdropper's channel. We also assume a quasi-static Rayleigh block fading model in this wiretap channel. Under this assumption, the channel coefficients in the main and the eavesdropper's channels are independent identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance across the coherence blocks. The channel coefficients hold constant within a coherence block but vary independently from block to block.

We denote h_i as the channel coefficient between the i-th transmit antenna at Alice and the received antenna at Bob, where $h_i \sim \mathcal{CN}(0,1)$, such that the main channel vector is expressed as $\mathbf{h} = [h_1, h_2, \cdots, h_M]$. The received symbol at Bob is written as

$$y = \mathbf{h}\mathbf{x} + n_b, \tag{1}$$

where n_b is the additive white Gaussian noise (AWGN) at Bob with variance σ_b^2, and $\mathbf{x}$ is the transmitted vector. Also, we denote g_i as the channel coefficient between the i-th transmit antenna at Alice and the received antenna at Eve, where $g_i \sim \mathcal{CN}(0,1)$. Correspondingly, the eavesdropper's channel vector is expressed as $\mathbf{g} = [g_1, g_2, \cdots, g_M]$. As such, the received symbol at Eve is written as

$$z = \mathbf{g}\mathbf{x} + n_e, \tag{2}$$

where n_e is the AWGN at Eve with variance σ_e^2.

In this work, we consider that Alice is able to acquire partial instantaneous knowledge about $\mathbf{h}$ sent from Bob through a reverse link. However, since Eve usually performs as an passive user (i.e., there is no reverse link between Alice and Eve), we consider that Alice cannot obtain any instantaneous knowledge about $\mathbf{g}$ from Eve.

A. Partial Knowledge of the Main Channel

Due to the constrained resource for Alice's CSI acquisition, Alice typically learns about the partial knowledge of the main channel through a finite amount of training and feedback overhead. To characterize the imperfect CSI at Alice, we propose a three-period block transmission protocol to capture the essential features of the practical FDD system. In our protocol, each block consists of a training period for a duration of τ_t, a feedback period for a duration of τ_f, and a data transmission period for a duration of τ_d. If the coherence block duration is τ_c, we have the constraint

$$\tau_c = \tau_t + \tau_f + \tau_d. \tag{3}$$

The scheduling period partition of a specific coherence block is depicted in Fig. 2. Specifically, the first M time slots (i.e., from T_1 to T_M) form the training period, the following time slot (i.e., T_{M+1}) forms the digital feedback period, and the last time slot (i.e., T_{M+2}) forms the data transmission period.

1) Channel Estimation: During the training period, the transmitter sequentially scans the transmit antennas to transmit T training symbols. To obtain a reliable estimation, the number of training symbols needs to be no less than the number of transmit antennas [26]. We assume that Bob computes the linear MMSE estimation of $\mathbf{h}$ given the received vector corresponding to T training symbols. Since the elements of $\mathbf{h}$ are assumed to be complex i.i.d Gaussian random variables, we have $\mathbf{h} = \hat{\mathbf{h}} + \mathbf{m}$, where the estimated channel $\hat{\mathbf{h}}$ and the estimated error $\mathbf{m}$ contain i.i.d complex Gaussian elements, satisfying $\hat{\mathbf{h}} \sim \mathcal{CN}\left(0, (1-\sigma_m^2)\mathbf{I}_M\right)$ and $\mathbf{m} \sim \mathcal{CN}\left(0, \sigma_m^2 \mathbf{I}_M\right)$, respectively. According to [29], the variance of the estimation error $\mathbf{m}$ can be written as

$$\sigma_m^2 = \frac{1}{1 + \rho_b T/M}, \tag{4}$$

where $\rho_b = P/\sigma_b^2$ denotes the average SNR of the received signal at Bob without estimation errors. Based on the estimated channel information $\hat{\mathbf{h}}$, Bob is able to obtain the channel direction information (CDI) of the main channel, i.e., $\mathbf{d} = \hat{\mathbf{h}}/\|\hat{\mathbf{h}}\|$.

2) CDI Quantization: After transmitting the training symbols, Alice waits for Bob to convey B bits over a feedback channel. The B bits specify the CDI associated with a particular beamforming vector. We clarify that Bob believes $\mathbf{d}$ as the accurate CDI of the main channel, based on which Bob selects the beamforming vector $\mathbf{f}$ from a quantization codebook $\mathcal{V}$ formed by 2^B unit-norm vectors, i.e., $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{2^B}\}$. Given $\mathcal{V}$, Bob selects the SNR-maximizing beamforming vector, which yields

$$\mathbf{f} = \arg\max_{\mathbf{v}_i \in \mathcal{V}} |\mathbf{d}\mathbf{v}_i|^2. \tag{5}$$

Bob then feeds back the index of the selected beamforming vector to Alice during time slot T_{M+1}.

The optimization of the codebook $\mathcal{V}$ has been thoroughly studied in [30, 31]. Albeit starting from different perspectives, [30] and [31] concluded the same criterion for beamformer design. In this criterion, a good codebook is the one which minimizes the maximum correlation between any pair of beamforming vectors. In this paper, we adopt the quantization codebook generated by this criterion. Under this criterion, we define $\cos^2\theta = |\mathbf{d}\mathbf{f}|^2$, the cumulative distribution function (CDF) of $\cos^2\theta$ is approximated as

$$F_{\cos^2\theta}(x) = \begin{cases} 0, & 0 \le x < 1-\varepsilon \\ 1 - 2^B(1-x)^{M-1}, & 1-\varepsilon \le x \le 1, \end{cases} \tag{6}$$

where $\varepsilon = 2^{-\frac{B}{M-1}}$ reflects the maximum error of the quantization at Bob.

B. Artificial-Noise-Aided Scheme

For the passive eavesdropping scenario, the instantaneous eavesdropper's channel knowledge is unknown to Alice, and in this case a popular method to guarantee secure communication is the artificial-noise-aided beamforming scheme adopted in [14–16, 23, 32], which is outlined as follows. After receiving the index of the selected beamformer $\mathbf{f}$, Alice generates a $M \times M$ precoding matrix as $\mathbf{W} = [\mathbf{f}, \mathbf{F}]$, where the columns of $\mathbf{F}$ form an orthonormal basis for the null space of $\mathbf{f}^\dagger$. The $M \times 1$ transmitted symbol vector $\mathbf{x}$ at Alice is given by $\mathbf{x} = \mathbf{f}u + \mathbf{F}\mathbf{v}$, where u denotes the information-bearing signal, and $\mathbf{v}$ represents the artificial noise. The variance of u is P_u, and the $M-1$ elements of $\mathbf{v}$ are i.i.d complex Gaussian random variables with zero mean and variance P_v. We consider a transmit power constraint, denoted by P, where $P = P_u + (M-1)P_v$. We denote the fraction of total power allocated to the information signal as ϕ. Hence, the power of the information-bearing signal and the artificial noise can be written as $P_u = P\phi$ and $P_v = P(1-\phi)/(M-1)$, respectively. In the following, we refer to $\kappa = \frac{P_u}{P_v} = \frac{M-1}{\phi^{-1}-1}$ as the power allocation ratio (PAR) between the information-bearing signal and the artificial noise.

By applying this artificial-noise-aided beamforming design, the received symbols at Bob and Eve become

$$y = \hat{\mathbf{h}}\mathbf{f}u + \hat{\mathbf{h}}\mathbf{F}\mathbf{v} + \mathbf{m}\mathbf{x} + n_b, \tag{7}$$

and

$$z = \mathbf{g}\mathbf{f}u + \mathbf{g}\mathbf{F}\mathbf{v} + n_e, \tag{8}$$

respectively. It is known that the perfect case corresponds to $\mathbf{f} = \hat{\mathbf{h}}^\dagger/\|\hat{\mathbf{h}}\|$, such that $\mathbf{F} = null(\hat{\mathbf{h}}/\|\hat{\mathbf{h}}\|)$ and $\hat{\mathbf{h}}\mathbf{F} = \mathbf{0}$. In other words, under the perfect case the artificial noise only confuses Eve without leaking into the main channel. However, as described in Section II-A, this ideal scenario is not realistic. Practically, since Alice only obtains Bob's quantized beamforming vector (i.e., $\mathbf{f} \neq \hat{\mathbf{h}}^\dagger/\|\hat{\mathbf{h}}\|$), the artificial noise that is originally intended to disrupt the Eve's reception may now leak into the main channel (i.e., $\hat{\mathbf{h}}\mathbf{F} \neq \mathbf{0}$), causing the so-called noise leakage problem [22].

C. Wiretap Codes Design

We now determine the wiretap code parameters to achieve a certain secrecy rate given by $R_s = R_b - R_e$. Specifically, we construct a parameter pair (R_b, R_e) for the wiretap codes [4], where R_b denotes the codeword transmission rate and R_e denotes the rate redundancy which provides secrecy against eavesdropping. In this work, we adopt the fixed-rate encoding strategy to reduce the complexity of practical application. That is, R_b and R_e hold constant over the transmission blocks. The key challenge in our considered wiretap channel is how to determine (R_b, R_e) in the passive eavesdropping scenario. In this work, we address this challenge as follows: We find the smallest possible rate redundancy R_e to guarantee the security level of the system requirements; We find the optimal R_b by formulating the closed-form expression for the secrecy throughput and maximizing this performance metric.

III. Statistical Characterization of the SNRs

In preparation for the design of secure transmission, in this section we focus on characterizing the statistics of the received SNRs at Bob and Eve. Specifically, by treating the combination of the estimation noise and the thermal noise at Bob as the worst-case Gaussian noise, we derive the statistics of Bob's received SNR. Also, by ignoring the estimation error and thermal noise at Eve for a robust design, we present the statistics of Eve's received SNR.

A. The Received SNR at Bob

Since Bob has no knowledge of his channel estimation error $\mathbf{m}$, he is unable to perfectly determine his instantaneous received SNR. Under this scenario, it is shown in [29] and [33] that treating $\mathbf{m}\mathbf{x} + n_b$ in (7) as a zero-mean Gaussian random variable $\hat{n}_b$ with variance $\sigma_{\hat{b}}^2 = \sigma_m^2 + \sigma_b^2$ minimizes the mutual information and therefore gives a lower bound on the capacity with channel estimation and finite feedback bits. By applying this method, we rewrite (7) as

$$y = \hat{\mathbf{h}}\mathbf{f}u + \hat{\mathbf{h}}\mathbf{F}\mathbf{v} + \hat{n}_b. \tag{9}$$

Based on (9), the actual instantaneous SNR of the main channel is given by

$$\gamma_b = \frac{\rho_u |\hat{\mathbf{h}}\mathbf{f}|^2}{\rho_v \|\hat{\mathbf{h}}\mathbf{F}\|^2 + 1} = \frac{\rho_u \|\hat{\mathbf{h}}\|^2 |\mathbf{d}\mathbf{f}|^2}{\rho_v \|\hat{\mathbf{h}}\|^2 \|\mathbf{d}\mathbf{F}\|^2 + 1}, \tag{10}$$

where $\rho_u = P_u/\sigma_{\hat{b}}^2$ and $\rho_v = P_v/\sigma_{\hat{b}}^2$. Due to the fact that $|\mathbf{df}|^2+\|\mathbf{dF}\|^2 = 1$, we substitute $|\mathbf{df}|^2 = \cos^2\theta$ and $\|\mathbf{dF}\|^2 = \sin^2\theta$ into (10) and rewrite γ_b as

$$\gamma_b = \frac{\rho_u \|\hat{\mathbf{h}}\|^2 \cos^2\theta}{\rho_v \|\hat{\mathbf{h}}\|^2 \sin^2\theta + 1}, \tag{11}$$

where $\|\hat{\mathbf{h}}\|^2$ has a Gamma distribution[2] with parameters $(M, 1-\sigma_m^2)$. For ease of notation, we define $\alpha = 1-\sigma_m^2$ such that $\|\hat{\mathbf{h}}\|^2 \sim Gamma(M,\alpha)$. In the following, we aim to derive the statistical distribution of γ_b by applying the method proposed in [34].

Lemma 1: We consider two independent Gamma random variables $X \sim Gamma(1,\alpha)$ and $Y \sim Gamma(M-1,\alpha)$ and define $U = \varepsilon Y$ and $V = X + (1-\varepsilon)Y$. We also define $I = \|\hat{\mathbf{h}}\|^2 \sin^2\theta$ and $S = \|\hat{\mathbf{h}}\|^2 \cos^2\theta$. Under the quantization cell approximation in (6), the joint distribution of (I,S) is the same as that of (U,V).

Proof: The proof is given in Appendix A. ∎

Based on ***Lemma 1***, we conclude that the interference term $\rho_v\|\hat{\mathbf{h}}\|^2 \sin^2\theta \sim Gamma(M-1,\rho_v\alpha\varepsilon)$, and the received signal power $\rho_u\|\hat{\mathbf{h}}\|^2\cos^2\theta \sim Gamma(1,\rho_u\alpha) + Gamma(M-1,\rho_u\alpha(1-\varepsilon))$. Note that the signal and interference are correlated through Y.

Lemma 2: Aided by the random variables defined in ***Lemma 1***, we rewrite γ_b as

$$\gamma_b = \frac{\rho_u S}{\rho_v I + 1} = \frac{\rho_u X + \rho_u(1-\varepsilon)Y}{\rho_v \varepsilon Y + 1}. \tag{14}$$

Then, we derive the CDF of γ_b as

$$F_{\gamma_b}(x) = \begin{cases} 1 - e^{-\frac{x}{\rho_u\alpha}}\beta^{M-1}, & x \geq \gamma_0 \\ F_Y(\lambda) - e^{-\frac{x}{\rho_u\alpha}}\beta^{M-1}F_\Upsilon(\lambda), & x < \gamma_0, \end{cases} \tag{15}$$

where $\beta = \frac{\kappa\varepsilon^{-1}}{\kappa+x}$, $\gamma_0 = \left(\varepsilon^{-1}-1\right)\kappa$, $\lambda = \frac{x}{\rho_u(1-\varepsilon)-\rho_v\varepsilon x}$ and Υ has a Gamma distribution $Gamma(M-1,\alpha\beta)$.

Proof: The proof is given in Appendix B. ∎

In this work, we refer to γ_b as the channel quality information (CQI) of the main channel. It is worth mentioning that at each transmission block, the instantaneous knowledge of γ_b is known at Bob. We state that Bob's knowing γ_b is significant for our following transmission scheme design in Section IV.

B. The Received SNR at Eve

Since the estimation error and the thermal noise at Eve are typically unknown to Alice, a robust approach needs to be designed for the worst-case scenario. That is, we assume that there is no thermal noise and no estimation error at Eve. As such, the received signal at Eve is simplified as

$$z = \mathbf{gf}u + \mathbf{gFv}. \tag{16}$$

By doing so, the actual instantaneous SNR of the eavesdropper's channel is given by

$$\gamma_e = \frac{|\mathbf{gf}|^2 P_u}{\|\mathbf{gF}\|^2 P_v} = \kappa\frac{|\mathbf{gf}|^2}{\|\mathbf{gF}\|^2}. \tag{17}$$

Since the unitary matrix $\mathbf{W} = [\mathbf{f},\mathbf{F}]$ is merely determined by the CDI of the main channel, we state that $\mathbf{W}$ is independent of $\mathbf{g}$. This leads to the fact that the statistics of $\mathbf{gW}$ is the same with $\mathbf{g}$. That is, the entires of $\mathbf{gW}$ are i.i.d. Gaussian random variables with zero mean and unit variance. Using [16, Eq. (5)], the CDF of γ_e is characterized as

$$F_{\gamma_e}(x) = 1 - \left(1 + \frac{x}{\kappa}\right)^{1-M}, \tag{18}$$

which is merely determined by the number of the transmit antenna M and the parameter κ. We clarify that (18) enables us to investigate the secrecy performance without knowing the average received SNR at Eve. The only required condition is the Rayleigh fading environment. As such, (18) is especially valid to the situation where there is no communication link between Alice and Eve, e.g., Eve is a completely silent and passive eavesdropper.

IV. Secure Transmission Design

In this section, we focus on designing the secure transmission scheme in the considered wiretap channel to guarantee the secrecy performance against eavesdropping. Specifically, we first characterize the secrecy outage performance and determine the smallest rate redundancy given a secrecy outage constraint. Then we adopt the on-off scheme to perform our transmission design and facilitate the closed-form expression for the secrecy throughput. Finally, we investigate the optimization problem of maximizing the secrecy throughput with fixed training and feedback overhead by proposing an iterative algorithm to determine the optimal parameters.

A. Secrecy Outage Probability

For the passive eavesdropping scenario, the instantaneous knowledge of γ_e is difficult to be unveiled at Alice, and thus the perfect secrecy is unavailable. In this scenario, the secrecy outage probability is usually used to characterize the security level from a probabilistic perspective [27, 35]. In particular, the secrecy outage event occurs when the designed rate redundancy R_e is lower than the instantaneous capacity of the eavesdropper's channel, i.e., $C_e = \log_2(1+\gamma_e)$. Mathematically, for a given rate redundancy R_e, the secrecy outage probability is formulated as

$$p_{so} = \Pr\{R_e < C_e\} = \Pr\left\{2^{R_e} - 1 < \gamma_e\right\}. \tag{19}$$

Using the CDF of γ_e in (18), we derive p_{so} as

$$p_{so} = \left(1 + \left(\frac{2^{R_e}-1}{\kappa}\right)\right)^{1-M}. \tag{20}$$

[2]A gamma distributed random X with parameters (M,ξ) is the sum of M i.i.d. exponential random variables with ξ mean. The statistical distributions of X are given by

$$f_X(x) = \frac{x^{M-1}e^{-x/\xi}}{\Gamma(M)\,\xi^M}, \tag{12}$$

$$F_X(x) = 1 - \sum_{m=0}^{M-1}\frac{(x/\xi)^m e^{-x/\xi}}{\Gamma(m+1)}. \tag{13}$$

To guarantee the security level of the system requirement, we impose a secrecy outage constraint δ in our transmission design. Under the secrecy outage constraint $p_{so} \leq \delta$, the minimum required rate redundancy is given by

$$R_{e,min} = \log_2(1+\gamma_{th}), \tag{21}$$

where $\gamma_{th} = (\delta^{\frac{1}{1-M}} - 1)\kappa$.

B. On-Off Transmission Scheme

To ensure that a positive secrecy rate R_s is available, the predetermined R_b needs to be always larger than $R_{e,min}$. For the convenience of following derivation, here, we denote R_b as $R_b = \log_2(1+\gamma_x)$, where γ_x is a constant parameter and needed to be optimally determined offline. The feasible region of γ_x is presented as follows: Firstly, $R_b > R_{e,min}$ implies the fact that the condition $\gamma_x > \gamma_{th}$ is needed. Secondly, we add another condition $\gamma_x \geq \gamma_0$ on γ_x for the following reasons. Firstly, this condition enables us to perform further mathematical analysis by using the CDF of γ_b given by (15). Secondly, we note that when B increases to some extent, the performance gain resulting from increasing B tends to be zero. Thus to efficiently take use of the feedback resource, the size of B should be restricted. We also note that the second condition $\gamma_x \geq \gamma_0$ becomes stricter as B increases. As such, the second condition guarantees that a good throughput performance can be obtained only when the predetermined B is not too large (e.g., $B < 10M$)[3].

In this work, we state that the optimization of γ_x should satisfy the above two conditions. Recall that $\gamma_0 = (2^{\frac{B}{M-1}} - 1)\kappa$ and $\gamma_{th} = (\delta^{\frac{1}{1-M}} - 1)\kappa$, and we present the combined condition on γ_x by defining two interesting cases.

- *Case 1 (Strong Security Requirement)*: $\delta \leq 2^{-B}$. Under this case, we have $\gamma_0 \leq \gamma_{th}$, such that we investigate the design of γ_x only when γ_x satisfies

$$\gamma_x > \gamma_{th}. \tag{22}$$

- *Case 2 (Weak Security Requirement)*: $\delta > 2^{-B}$. Under this case, we have $\gamma_0 > \gamma_{th}$, such that we investigate the design of γ_x only when γ_x satisfies

$$\gamma_x \geq \gamma_0. \tag{23}$$

From the design perspective, (22) and (23) imply that the "on-off" transmission scheme with a threshold of γ_x can be adopted to perform our transmission design. Specifically, the secure transmission is allowed only when the CQI known at Bob (i.e., γ_b) is larger than γ_x. Using the CDF of γ_b in (15), the transmission probability is mathematically derived as

$$p_{tx} = \Pr\{\gamma_b \geq \gamma_x\} = \frac{2^B \kappa^{M-1} e^{-\frac{\gamma_x}{\rho_u \alpha}}}{(\kappa+\gamma_x)^{M-1}}. \tag{24}$$

It is worth mentioning that in addition to the B feedback bits for CDI quantization, at each transmission block Bob also needs to convey back an extra bit identifying the on/off state of the transmission. However, this 1-bit overhead is quite small compared with the B bits for CDI quantization. Thus in this work we omit this additional 1-bit to concentrate on the impact of CDI quantization on the secrecy performance.

C. Secrecy Throughput Maximization

The aim of our transmission design is to achieve the maximal secrecy throughput under the given secrecy outage constraint. Here, the secrecy throughput is defined as [36]

$$\eta = p_{tx}(1-\delta)(R_b - R_{e,min}). \tag{25}$$

Notably, the secrecy throughput measures the average rate of the transmitted message which is kept confidential to the eavesdropper. Substituting (24) into (25), we formulate η as

$$\eta = \frac{2^B(1-\delta)e^{-\frac{\gamma_x}{\rho_u\alpha}}}{(\kappa+\gamma_x)^{M-1}\kappa^{1-M}} \log_2\left(\frac{1+\gamma_x}{1+\gamma_{th}}\right). \tag{26}$$

To make our following analysis clear, here, we normalize the noise power at Bob to 1, i.e., $\sigma_{\hat{b}}^2 = 1$. By doing so, we have $\sigma_{\hat{b}}^2 = 1+\sigma_m^2 = 2-(1-\sigma_m^2) = 2-\alpha$. As such, $\rho_u = P_u/\sigma_{\hat{b}}^2$ can be rewritten as $\rho_u = P\phi/(2-\alpha)$. Moreover, since $\kappa = (M-1)/(\phi^{-1}-1)$, we find that $\phi = \kappa/(M-1+\kappa)$ and the feasible region of κ is $[0,\infty)$. Based on these transformations, we rewrite (26) as

$$\eta = \frac{2^B(1-\delta)e^{-\frac{\gamma_x(2-\alpha)(M-1+\kappa)}{P\alpha\kappa}}}{(\kappa+\gamma_x)^{M-1}\kappa^{1-M}} \log_2\left(\frac{1+\gamma_x}{1+h(\delta)\kappa}\right), \tag{27}$$

where $h(\delta) = \delta^{\frac{1}{1-M}} - 1$. For the given training and feedback overhead (i.e., T and B are fixed), we find that η is merely a function of κ and γ_x. This motivates us to address the following optimization problem:

Problem 1: Given the fixed T and B, what are the optimal κ and γ_x that maximize the secrecy throughput under a given secrecy outage constraint? This problem is mathematically expressed as

$$\begin{aligned} &\max_{\kappa,\gamma_x}\ \eta, \\ &\text{s.t.}\quad \kappa \geq 1, \gamma_x \geq \gamma_0, \gamma_x > \gamma_{th}, \end{aligned} \tag{28}$$

where η is expressed in (27). In the following, we are going to calculate the optimal values of κ and γ_x.

Due to the interaction between κ and γ_x, the above optimization problem is not directly intractable. As such, we divide this joint optimization problem into two subproblems and provide an iterative and mutual optimization algorithm to handle it. In particular, we optimize κ and γ_x iteratively according to the following sequence: $\kappa[0] \to \gamma_x[0] \to \cdots \to \kappa[n-1] \to \gamma_x[n-1] \to \kappa[n] \to \gamma_x[n] \to \cdots \to \kappa_{opt} \to \gamma_{x,opt}$. This iterative process starts by initializing κ and γ_x. For each iteration step n, we calculate the optimal $\kappa[n]$ for a given $\gamma_x[n-1]$ inherited from the last iteration, and then we calculate the optimal $\gamma_x[n]$ with the fixed $\kappa[n]$. We repeat this process until no further improvement on the secrecy throughput can be obtained. Next, we present the two subproblems with their solutions in the following.

[3]It is worth mentioning that since B is not large, the quantization error is not very small. Under this case, if we set $P_u < P_v$ (i.e., $\kappa < 1$), the noise leakage problem is likely to grow far worse, seriously degrading the secrecy performance. As such, in this work we only investigate the case of $\kappa \geq 1$.

1) Optimal κ for a Fixed γ_x: By taking the first-order derivative of η in (27) on κ, we derive $\frac{\partial \eta}{\partial \kappa}$ as

$$\frac{\partial \eta}{\partial \kappa} = \frac{2^B (1-\delta)}{\ln 2} \zeta(\kappa) J(\kappa), \tag{29}$$

where $\zeta(\kappa) = \frac{1}{\kappa^2} e^{-\frac{\gamma_x}{\rho_u \alpha}} (\frac{\kappa}{\kappa+\gamma_x})^{M-1} \ln\left(\frac{1+\gamma_x}{1+h(\delta)\kappa}\right)$ and $J(\kappa)$ is given by

$$J(\kappa) = \frac{(M-1)\gamma_x \kappa}{\gamma_x + \kappa} - \frac{h(\delta)\kappa^2}{1+h(\delta)\kappa} \ln^{-1}\left(\frac{1+\gamma_x}{1+h(\delta)\kappa}\right) + \frac{\gamma_x (2-\alpha)(M-1)}{P\alpha}. \tag{30}$$

Since $\zeta(\kappa)$ is always positive, the sign of $\frac{\partial \eta}{\partial \kappa}$ follows that of $J(\kappa)$. As such, we are able to justify the monotonicity of $\eta(\kappa)$ by analyzing $J(\kappa)$. However, since $J(\kappa)$ in (30) is still complicated, we further take the first-order derivative of $J(\kappa)$ and derive $J^{'}(\kappa)$ as

$$J^{'}(\kappa) = \frac{(M-1)\gamma_x^2}{(\gamma_x+\kappa)^2} - \left(\frac{1-\left(\frac{1}{1+h(\delta)\kappa}\right)^2}{\ln\left(\frac{1+\gamma_x}{1+h(\delta)\kappa}\right)} + \frac{\left(1-\frac{1}{1+h(\delta)\kappa}\right)^2}{\ln^2\left(\frac{1+\gamma_x}{1+h(\delta)\kappa}\right)} \right). \tag{31}$$

We find from (31) that $J^{'}(\kappa)$ is a decreasing function of κ. Next, we discuss the optimal κ maximizing $\eta(\kappa)$ for the following two cases.

- *Case 1*: $\delta \leq 2^{-B}$. This case is referred to as *Strong Security Requirement* in Section IV-B, under which the transmission condition is $\gamma_{th} < \gamma_x$, such that the feasible region of κ is $[0, \frac{\gamma_x}{h(\delta)})$. Using (31) and (30), it is easy to obtain that $J^{'}(0) > 0$, $J^{'}(\frac{\gamma_x}{h(\delta)}) < 0$, $J(0) > 0$ and $J(\frac{\gamma_x}{h(\delta)}) < 0$. That is, $J(\kappa)$ first increases from a positive value then decreases to a negative value as κ increases. As such, we conclude that $\eta(\kappa)$ first increases and then decreases as κ increases. And the optimal κ maximizing $\eta(\kappa)$ is the unique root of $J(\kappa) = 0$.
- *Case 2*: $\delta > 2^{-B}$. This case is referred to as *Weak Security Requirement* in Section IV-B, under which the transmission condition is $\gamma_0 \leq \gamma_x$, such that the feasible region of κ is $[0, \frac{\gamma_x}{\varsigma}]$, where $\varsigma = 2^{\frac{B}{M-1}} - 1$. The optimal κ that maximizes $\eta(\kappa)$ depends on the sign of $J^{'}(\frac{\gamma_x}{\varsigma})$ and $J(\frac{\gamma_x}{\varsigma})$: 1) If $J^{'}(\frac{\gamma_x}{\varsigma}) \geq 0$, the optimal κ is $\frac{\gamma_x}{\varsigma}$; 2) If $J^{'}(\frac{\gamma_x}{\varsigma}) < 0$ and $J(\frac{\gamma_x}{\varsigma}) \geq 0$, the optimal κ is $\frac{\gamma_x}{\varsigma}$; 3) If $J^{'}(\frac{\gamma_x}{\varsigma}) < 0$ and $J(\frac{\gamma_x}{\varsigma}) < 0$, the optimal κ is the unique root of $J(\kappa) = 0$.

Based on the above discussions, we present the optimal γ_x maximizing η with a fixed κ in the following theorem.

***Theorem** 1:* For a fixed γ_x, the optimal κ maximizing η is

$$\kappa^* = \max\{\kappa^\circ, 1\}, \tag{32}$$

where

$$\kappa^\circ = \begin{cases} \kappa^\star, & \delta \leq 2^{-B}, \\ \kappa^\star, & \delta > 2^{-B}, J^{'}(\frac{\gamma_x}{\varsigma}) < 0, J(\frac{\gamma_x}{\varsigma}) < 0 \\ \frac{\gamma_x}{\varsigma}, & \delta > 2^{-B}, J^{'}(\frac{\gamma_x}{\varsigma}) < 0, J(\frac{\gamma_x}{\varsigma}) \geq 0 \\ \frac{\gamma_x}{\varsigma}, & \delta > 2^{-B}, J^{'}(\frac{\gamma_x}{\varsigma}) \geq 0, \end{cases} \tag{33}$$

and $\kappa^\star$ satisfies

$$0 = \frac{(M-1)\gamma_x \kappa^\star}{\gamma_x + \kappa^\star} - \frac{h(\delta)(\kappa^\star)^2}{1+h(\delta)\kappa^\star} \ln^{-1}\left(\frac{1+\gamma_x}{1+h(\delta)\kappa^\star}\right) + \frac{\gamma_x (2-\alpha)(M-1)}{P\alpha}. \tag{34}$$

We state that although we are not able to derive an explicit expression for $\kappa^\star$, we can efficiently calculate it through the binary search method.

2) Optimal γ_x for a Fixed κ: By taking the first-order derivative of η on γ_x, we derive $\frac{\partial \eta}{\partial \gamma_x}$ as

$$\frac{\partial \eta}{\partial \gamma_x} = \frac{2^B (1-\delta)}{\ln 2} \frac{\kappa^{M-1} e^{-\frac{\gamma_x}{\rho_u \alpha}}}{(\kappa+\gamma_x)^M} K(\gamma_x), \tag{35}$$

where

$$K(\gamma_x) = \frac{\kappa+\gamma_x}{1+\gamma_x} - \left(M - 1 + \frac{\kappa+\gamma_x}{\rho_u \alpha}\right) \ln\left(\frac{1+\gamma_x}{1+\gamma_{th}}\right). \tag{36}$$

It is easy to find that the sign of $\frac{\partial \eta}{\partial \gamma_x}$ follows that of $K(\gamma_x)$. In other words, we can examine the monotonicity of $\eta(\gamma_x)$ by examine the sign of $K(\gamma_x)$. We find from (36) that $\lim_{\gamma_x \to \infty} K(\gamma_x) = -\infty$ and $K(\gamma_x)$ is a decreasing function of γ_x when $\kappa \geq 1$ is satisfied. Next, we discuss the optimal γ_x maximizing $\eta(\gamma_x)$ for the following two cases.

- *Case 1*: $\delta \leq 2^{-B}$. Under this case, the feasible region of γ_x is (γ_{th}, ∞). Since $K(\gamma_{th}) > 0$, we conclude that $K(\gamma_x)$ becomes first positive and then negative as γ_x increases. As such, the optimal γ_x maximizing $\eta(\gamma_x)$ is the unique root of $K(\gamma_x) = 0$.
- *Case 2*: $\delta > 2^{-B}$. Under this case, the feasible region of γ_x is $[\gamma_0, \infty)$, and the monotonicity of $\eta(\gamma_x)$ depends on the sign of $K(\gamma_0)$. In particular, if $K(\gamma_0) \leq 0$, $\eta(\gamma_x)$ monotonically decreases with γ_x, such that the maximum is achieved at $\gamma_x = \gamma_0$. However, if $K(\gamma_0) > 0$, $\eta(\gamma_x)$ first increases and then decreases with γ_x, and the maximum is achieved at the unique root of $K(\gamma_x) = 0$.

Now, we present the optimal γ_x maximizing η with a fixed κ in the following theorem.

***Theorem** 2:* For a fixed κ, the optimal γ_x maximizing η is

$$\gamma_x^* = \begin{cases} \gamma^\star, & \delta \leq 2^{-B}, \\ \gamma^\star, & \delta > 2^{-B}, K(\gamma_0) > 0, \\ \gamma_0, & \delta > 2^{-B}, K(\gamma_0) \leq 0, \end{cases} \tag{37}$$

where $\gamma^\star$ satisfies

$$\frac{\kappa+\gamma^\star}{1+\gamma^\star} - \left(M - 1 + \frac{\kappa+\gamma^\star}{\rho_u \alpha}\right) \ln\left(\frac{1+\gamma^\star}{1+\gamma_{th}}\right) = 0. \tag{38}$$

Also, the explicit expression for $\gamma^\star$ is difficult to derive, and we still adopt the binary search method to calculate it.

By summarizing the above analysis to the two subproblems, we present the detailed process to obtain the solutions for ***Problem** 1* in **Algorithm 1**. As indicated in [26], this two-step iterative algorithm can converge fast to a stationary point of the optimization problem in (28). The proof is not provided in this article, and readers may refer to [26] for detailed process. Here, we highlight that this algorithm is an efficient method to deal with ***Problem** 1*, which can be observed in Section VI.

Algorithm 1 Proposed Iterative Algorithm for Solving the Problem in (28).

1: Initialization: $n = 1$, $\kappa[1] = 1$, $\gamma_x[1] = 1$, $\eta(\kappa[0], \gamma_x[0]) = 0$ and $\epsilon > 0$.
2: Calculate $\eta(\kappa[1], \gamma_x[1])$.
3: **While** $|\eta(\kappa[n], \gamma_x[n]) - \eta(\kappa[n-1], \gamma_x[n-1])| \geq \epsilon$ **do**
4: $n = n + 1$;
5: Given $\gamma_x[n-1]$, find the optimal $\kappa[n]$ using (32);
6: Given $\kappa[n]$, find the optimal $\gamma_x[n]$ using (37);
7: Calculate the secrecy throughput $\eta(\kappa[n], \gamma_x[n])$;
8: **end while**
9: Output: $\kappa[n]$ and $\gamma_x[n]$.

V. Net Secrecy Throughput

In this section, we focus on a new secrecy performance metric, referred to as the NST. We attribute this metric to the non-trivial tradeoff between CSI acquisition and data transmission in block fading channels with a fixed coherence time. In such channels, a large amount of training and feedback overhead improves the CSI quality but reduces the time spent on data transmission, while a small amount of training and feedback overhead leads to a poor CSI quality, which in turn decreases the achievable secrecy throughput. Motivated by this fact, in this section we adopt the NST as the primary secrecy performance metric and determine the optimum training and feedback overhead maximizing the NST.

A. NST as a New Secrecy Performance Metric

Recall the analysis in Section II-A, the signaling overhead for channel acquisition (i.e., T and B) consumes $\tau_t + \tau_f$ time resource and only $\tau_c - \tau_t - \tau_f$ resource is used to transmit D data symbols. Based on this signaling overhead model, we define the NST as a closed-form expression given by

$$\eta_{\text{net}} = \left(1 - \frac{\tau_t + \tau_f}{\tau_c}\right)\eta, \tag{39}$$

where η is the secrecy throughput defined in (25). Here, we assume that both Alice and Bob use the same rate (i.e., r) to transmit the forward and feedback signals, such that $\tau_t = T/r$ and $\tau_f = \mu B/r$, where μ is a conversion factor relating bits to symbols. Substituting τ_t and τ_f into (39), we rewrite the NST as

$$\eta_{\text{net}} = \left(1 - \frac{T + \mu B}{L}\right)\eta, \tag{40}$$

where $L = r\tau_c$ is the symbol size of each coherence block.

B. NST Maximization

Apart from the closed-form expression for the NST indicated by (40), one may also want to know the optimal system parameters that maximize the NST for a given block size. This motivates us to investigate the following optimization problem:

Problem 2: Given the coherence block length L, what are the optimal T, B, κ and γ_x that maximize the NST under a given secrecy outage constraint? This problem is mathematically expressed as

$$\begin{aligned} \max_{T,B,\kappa,\gamma_x} \quad & \eta_{\text{net}}, \\ \text{s.t.} \quad & T + \mu B + D = L, \\ & \kappa \geq 1, \gamma_x \geq \gamma_0, \gamma_x > \gamma_{th}, \end{aligned} \tag{41}$$

where η_{net} is expressed in (40).

Recall that in Section IV-C, we have investigated the ***Problem 1*** and presented the algorithm to find the optimal κ and γ_x maximizing the secrecy throughput under the given T and B. That is, if we fix the values of T and B, we are able to obtain the optimal κ and γ_x maximizing η by applying **Algorithm 1**. Moreover, we find from (40) that for fixed T and B, the optimal κ and γ_x maximizing η_{net} are equivalent to the optimal κ and γ_x maximizing η. As such, ***Problem 2*** can be transformed to the following optimization problem.

Problem 3: Given the coherence block length L, what are the optimal T and B that maximize the NST under a given secrecy outage constraint? This problem is mathematically expressed as

$$\begin{aligned} \max_{T,B} \quad & \left(1 - \frac{T + \mu B}{L}\right)\eta\left(\kappa_{opt}, \gamma_{x,opt}\right), \\ \text{s.t.} \quad & T + \mu B + D = L, \end{aligned} \tag{42}$$

where κ_{opt} and $\gamma_{x,opt}$ are the solutions of ***Problem 1***. It is worth to mentioning that the values of κ_{opt} and $\gamma_{x,opt}$ vary with the change of T and B. That is, κ_{opt} and $\gamma_{x,opt}$ are the functions of T and B, but these two function relationships (i.e., $\kappa_{opt}(T, B)$ and $\gamma_{x,opt}(T, B)$) are implicit.

We clarify that ***Problem 3*** is a typical non-linear integer programming (NLIP) problem. The implicit relationships between $(\kappa_{opt}, \gamma_{x,opt})$ and (T, B) substantially increase the difficulties in solving this problem. According to the literature, there are only a few effective algorithms to practically solve it. The most popular ones are the heuristic search algorithm and the explicit enumeration method (EEM) [37]. The heuristic search algorithm is particularly useful to produce sub-optimal solutions to large-scale NLIP problems. However, it lacks the ability to obtain globally optimal solutions. Differing from the heuristic search algorithm, the EEM is known for its implementation simplicity in solving small-scale NLIP problems. Specifically, it can reliably provide the global optimal solutions to such problems with acceptable complexity.

Fortunately, we find that there are only two variables in the solution space of our optimization problem given by (42). This implies that the scale of our target problem is quite small. As such, we directly apply the easy-to-implement EEM to solve (42) in an efficient way. As per the rules of the EEM, we detail the process of obtaining the EEM solutions for (42) in **Algorithm 2**.

It is worth mentioning that the complexity of **Algorithm 2** is $O(L^2)$. This means that its complexity is in general acceptable since the typical block length is not very large. We also point out that the efficiency of **Algorithm 2** can be greatly improved by reducing the search space. Specifically, when $B > 10M$, the transmission threshold becomes very large (i.e., p_{tx} is quite

Algorithm 2 Proposed EEM for Solving the Problem in (42).

1: Initialization: $\eta_{\text{net,max}} = 0$, $T_{\text{opt}} = 0$ and $B_{\text{opt}} = 0$.
2: Global search:
3: **for** $T = \{M, M+1, \cdots, (L-1)\}$ **do**
4: **for** $B = \{M, M+1, \cdots, (L-T)/\mu\}$ **do**
5: Find κ_{opt} and $\gamma_{x,opt}$ by applying **Algorithm 1**.
6: Compute $\eta_{\text{net}} = \left(1 - \frac{T+\mu B}{L}\right) \eta\left(\kappa_{opt}, \gamma_{x,opt}\right)$.
7: **if** $\eta_{\text{net}} > \eta_{\text{net,max}}$ **do**
8: Update $\eta_{\text{net,max}} = \eta_{\text{net}}$, $T_{\text{opt}} = T$ and $B_{\text{opt}} = B$.
9: **end**
10: **end**
11: **end**
12: Output: T_{opt} and B_{opt}.

small), leading to the fact that the secrecy throughput is almost zero. Moreover, if the length of signaling overhead accounts for more than half of the whole block, the secrecy throughput performance will be degraded seriously.

VI. Numerical Results

In this section, we present some simulation results to characterize the secrecy performance of our designed transmission scheme. Firstly, we validate the accuracy level of our theoretical analysis with Mento Carlo simulations. Then we examine the secrecy throughput performance and the convergence of **Algorithm 1**. Finally, the optimal signaling overhead that maximizes the net secrecy throughput is characterized.

A. Verification of Theoretical Accuracy

In Section II-A, the approximated CDF of the quantization error was introduced for our theoretical analysis. This indicates that the derived expressions in our work are approximated results. Therefore, we first need to characterize the accuracy level of this approximation.

Fig. 3 plots the approximated secrecy throughput versus P for different values of B, along which the numerical secrecy throughput is provided using Monte Carlo simulation. The numerical points marked by '$*$' are averaged over 10,000 channel trials, and the quantization codebook is generated based on the design criterion in [30, 31]. It is seen from this figure that the difference between the approximated η and the numerical η is extremely minor. This indicates that the quantization approximation has an almost negligible effect on the secrecy throughput, such that the closed-form expression derived in (26) can accurately predict the secrecy throughput with limited training and feedback. Since the derivations in this work are based on the expression of secrecy throughput in (26), we state that Fig. 3 verifies the accuracy of our theoretical analysis in this work.

B. The Optimal κ and γ_x Maximizing Secrecy Throughput

In this subsection, we first examine the optimal κ and γ_x that maximize the secrecy throughput under fixed T and B by applying **Algorithm 1**, and then we study the convergence of this algorithm.

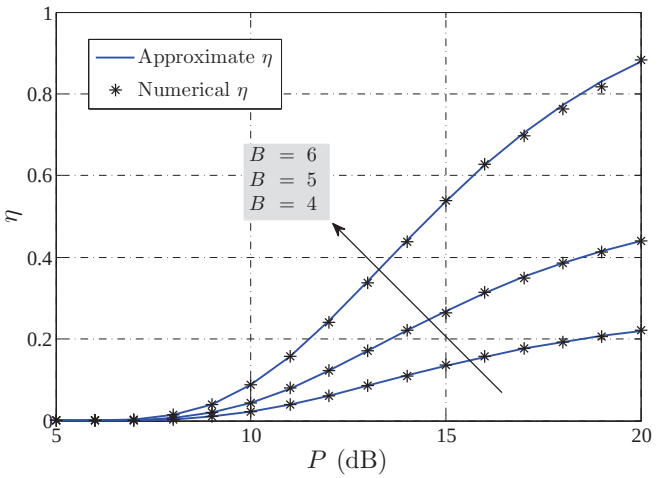

Fig. 3. Secrecy throughput versus P for different values of B with $M = 4$, $T = 4$, $\delta = 0.1$, $\kappa = 4$ and $\gamma_x = 15$.

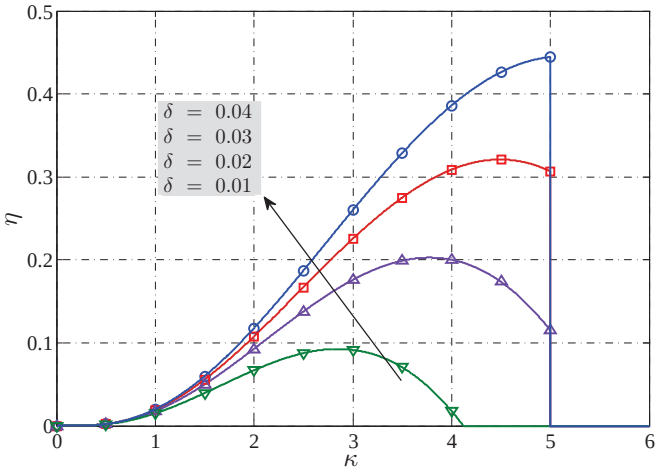

Fig. 4. Secrecy throughput versus κ for different values of δ with $M = 4$, $P = 100$, $T = 4$, $B = 6$ and $\gamma_x = 15$.

Fig. 4 plots the secrecy throughput versus κ for different values of δ with a fixed $\gamma_x = 15$. We first observe that strengthening the secrecy outage constraint (i.e., reducing δ) decreases the secrecy throughput. We further observe that the optimal κ maximizing η depends on the value of δ. In particular, if δ is small (e.g., $\delta = 0.01, 0.02, 0.03$), η first increases and then decreases as κ increases, and the optimal κ is the solution of $\partial\eta/\partial\kappa = 0$; while if δ becomes larger (e.g., $\delta = 0.04$), η monotonically increases with the increasing of κ, such that the optimal κ is at the right boundary.

Fig. 5 plots the secrecy throughput versus γ_x for different values of δ with a fixed $\kappa = 4$. In this figure, we observe that the optimal γ_x maximizing η depends on the value of δ. Specifically, if δ is small (e.g., $\delta = 0.01, 0.02, 0.03$), η first increases and then decreases as γ_x increases, and the optimal γ_x is the solution of $\partial\eta/\partial\gamma_x = 0$; while if δ is large (e.g., $\delta = 0.04$), η monotonically decreases with the increasing of γ_x, such that the optimal γ_x is at the left boundary. We highlight that the observations in Fig. 4 and 5 verify the correctness of

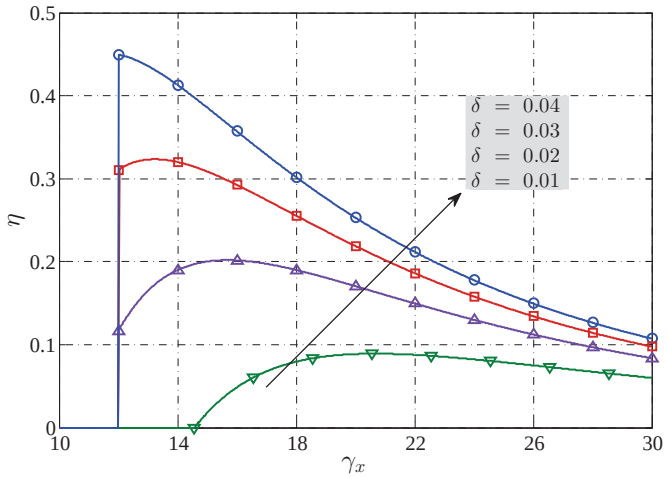

Fig. 5. Secrecy throughput versus γ_x for different values of δ with $M = 4$, $P = 100$, $T = 4$, $B = 6$ and $\kappa = 4$.

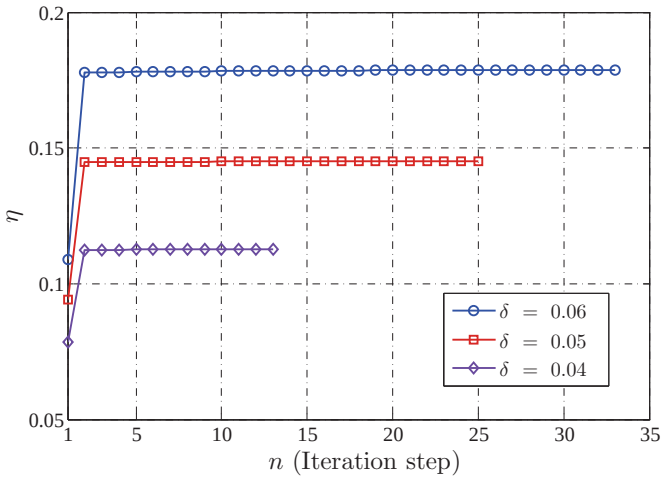

Fig. 7. The convergence of Algorithm 1 with $M = 4$, $P = 100$, $T = 4$, $B = 4$ and $\epsilon = 10^{-6}$.

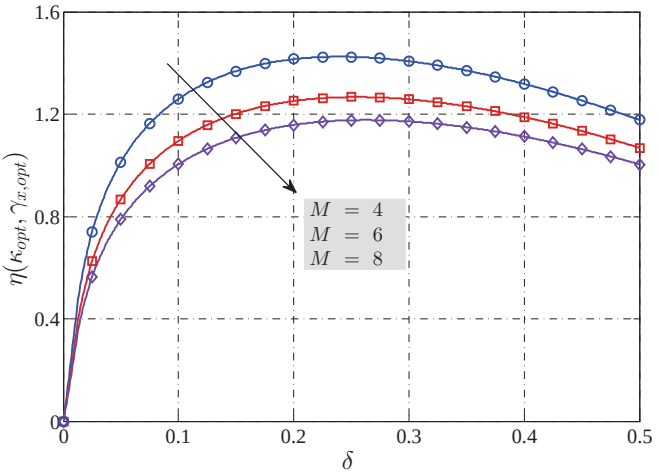

Fig. 6. Achievable secrecy throughput versus δ for different values of M with $P = 100$, $T = 8$, $B = 8$ and $\delta = 0.01$.

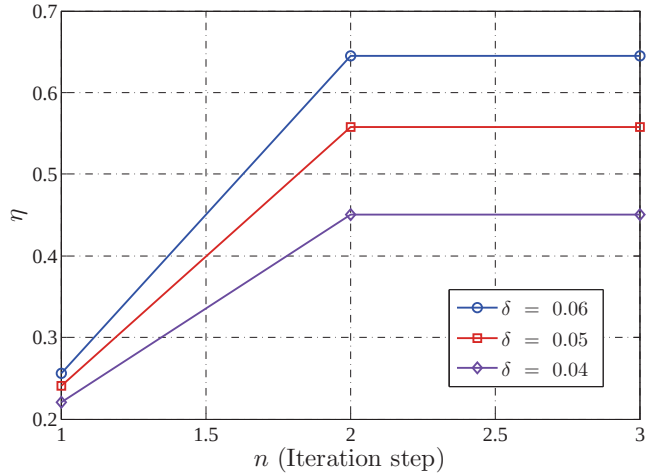

Fig. 8. The convergence of Algorithm 1 with $M = 4$, $P = 100$, $T = 4$, $B = 6$ and $\epsilon = 10^{-6}$.

our analysis in Section IV-C.

Fig. 6 plots the achievable secrecy throughput versus δ for different values of M with fixed T and B. The achievable secrecy throughput is achieved by applying **Algorithm 1** to determine the optimal κ and γ_x. We first observe that for a fixed M, the achievable secrecy throughput first increases then decreases as δ increases. That is, allowing for a certain degree of secrecy outage can indeed obtain an larger secrecy throughput, but bigger outage is not always better. We further observe that for a fixed δ, increasing the number of transmit antenna reduces the achievable secrecy throughput. This is not surprising since when M increases, the number of the training symbols and feedback bits for each transmit antenna decreases, such that the CSI quality obtained at Bob becomes worse and the secrecy performance is degraded.

Fig. 7 and 8 plot the convergence of Algorithm 1 with $\epsilon = 10^{-6}$ for two kinds of cases. In particular, Fig. 7 is for the case of *Strong Security Requirement* (i.e., $\delta \leq 2^{-B}$), and Fig. 8 is for the case of *Weak Secrecy Requirement* (i.e., $\delta > 2^{-B}$). We observe that when $\delta \leq 2^{-B}$, the convergence speed of Algorithm 1 depends on the value of δ. For example, when $\delta = 0.04$, Algorithm 1 converges under 13 iterative steps; While when $\delta = 0.06$, this number is 33. Although the needed iterative steps are different, they are generally smaller than 50 and thus the convergence speed is acceptable. On the other hand, when $\delta > 2^{-B}$, Algorithm 1 converges to a stationary point within only 3 iterative steps. This is due to the fact under this case, the boundary points are the optimal solutions, such that the algorithm can be converged quickly. These observations highlight that Algorithm 1 is an efficient algorithm to solve ***Problem 1***.

C. The Optimal Fraction of Overhead Maximizing NST

In this subsection, we numerically characterize the optimal fraction of the signaling overhead, $(T_{opt} + \mu B_{opt})/L$, maximizing the NST. To clarify, in the following we only consider BPSK feedback at Bob such that we have $\mu = 1$ [29].

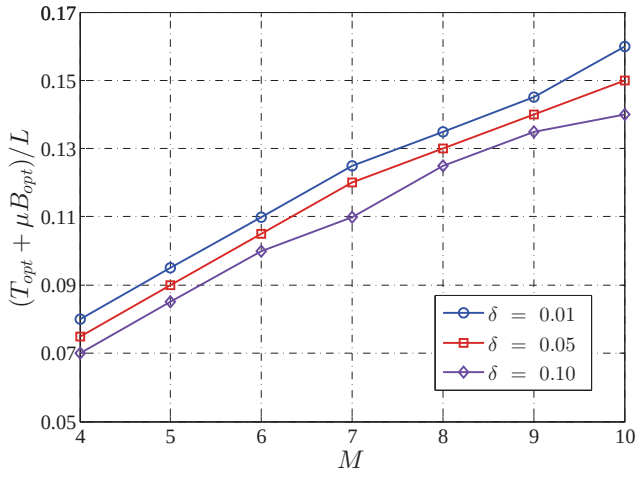

Fig. 9. Optimal overhead fraction versus M for different values of δ with $P = 100$ and $L = 200$.

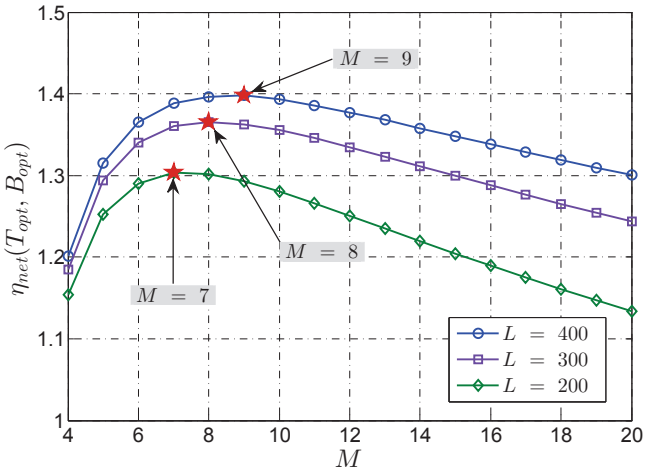

Fig. 10. Achievable net secrecy throughput versus M for different values of L with $P = 100$ and $\delta = 0.05$.

Fig. 9 plots the optimal fraction of the signaling overhead versus M for different values of δ. In this figure, the optimal signaling overhead is obtained by applying **Algorithm 2**. We first observe that the stronger the secrecy outage constraint, the more signaling overhead is needed. We further observe that the optimal overhead fraction is approximately proportional to M. This phenomenon can be explained as follows: To ensure Bob to acquire a satisfactory CSI quality, the signaling overhead averaged on each transmit antenna should not be reduced. Thus when M increases, the total signaling overhead should increase in proportion to M. Moreover, we find that this proportional factor is insensitive to the changes of δ.

Fig. 10 plots the achievable NST versus M for different values of L. Firstly, we observe that the achievable NST first increases then decreases as M increases. This phenomenon demonstrates that for a given L, the number of transmit antennas is not the more the better. Instead, there exists an optimal M that maximizes the NST. We also observe that the optimal M shifts to right when L increases. For example, we find that when L increases from 200 to 400, the optimal M grows from 7 to 9. Moreover, we observe that increasing L would increase the achievable NST, as one may expect. This figure highlights that the number of actively used transmit antennas needs to be determined according to the length of the coherence block. In particular, when there are plenty of transmit antennas available at Alice, we can merely exploit a a suitable number of them to achieve the maximum NST.

VII. Conclusions

In this paper, we investigated the design of secure transmission scheme in the MISOSE wiretap channel with limited training and feedback overhead. Assuming that Alice has the partial knowledge of the main channel but no knowledge of the eavesdropper's channel, we designed an artificial-noise-aided secure on-off transmission scheme under the secrecy outage constraint. Then we studied the problem of how to obtain the maximum secrecy throughput under fixed training and feedback overhead, based on which we further studied the problem how to optimally allocate the block resource to the signaling overhead. We numerically found that it is not always best that using all transmit antennas to perform secure transmission under limited signaling overhead.

To extend this work, it is of interest to consider the adaptive-rate encoding strategy for secure transmission design with limited signaling overhead. Further, in this work the scenario where Eve is equipped with multiple antennas has not been taken into consideration, which is also an important research direction of our future work.

Appendix A
Proof of Lemma 1

Firstly, we derive the joint distribution of (I, S), which is related to that of $(\|\hat{\mathbf{h}}\|^2, \sin^2\theta)$ by the transformation

$$f_{I,S}(u,v) = \|J_1\| f_{\|\hat{\mathbf{h}}\|^2,\sin^2\theta}(r,w). \tag{43}$$

Since $u = rw$ and $v = r(1-w)$, we have $r = u+v$ and $w = \frac{u}{u+v}$. Thus the Jacobian J_1 is given by

$$J_1 = \frac{\partial(r,w)}{\partial(u,v)} = \begin{pmatrix} 1 & 1 \\ \frac{v}{(u+v)^2} & \frac{-u}{(u+v)^2} \end{pmatrix}, \tag{44}$$

and $\|J_1\| = \frac{1}{u+v}$. Since $\|\hat{\mathbf{h}}\|$ and $\mathbf{d}$ are independent with each other, it follows that $\|\hat{\mathbf{h}}\|^2$ and $\sin^2\theta$ are independent. As such, the joint distribution of (I, S) is formulated as

$$f_{I,S}(u,v) = \frac{1}{u+v} f_{\|\hat{\mathbf{h}}\|^2}(u+v) f_{\sin^2\theta}\left(\frac{u}{u+v}\right), \tag{45}$$

Note that $\|\hat{\mathbf{h}}\|^2 \sim Gamma(M,\alpha)$ and the CDF of $\sin^2\theta$ is given by

$$F_{\sin^2\theta}(x) = \begin{cases} 2^B x^{M-1}, & 0 \le x \le \varepsilon \\ 1, & \varepsilon \le x \le 1. \end{cases} \tag{46}$$

As such, we have

$$\begin{aligned} & f_{I,S}(u,v) \\ &= \frac{1}{u+v} f_{\|\hat{\mathbf{h}}\|^2}(u+v)\, f_{\sin^2\theta}\left(\frac{u}{u+v}\right) \\ &= \frac{1}{u+v}\frac{(u+v)^{M-1}e^{-\frac{u+v}{\alpha}}}{\Gamma(M)\,\alpha^M} 2^B (M-1)\left(\frac{u}{u+v}\right)^{M-2} \\ &= \frac{2^B u^{M-2}}{\Gamma(M-1)\,\alpha^M} e^{-\frac{u+v}{\alpha}}, 0 \le \frac{u}{u+v} \le \varepsilon. \end{aligned} \tag{47}$$

Now consider two independent Gamma random variables $X \sim Gamma(1,\alpha)$ and $Y \sim Gamma(M-1,\alpha)$ and define $U = \varepsilon Y$ and $V = X + (1-\varepsilon)Y$. We would like to show that the joint distribution (U,V) is also given by (47).

In particular, the joint distribution of (U,V) is related to that of (X,Y) by the transformation

$$f_{U,V}(u,v) = \|J_2\| f_{X,Y}(x,y). \tag{48}$$

Since $u = \varepsilon y$ and $v = x + (1-\varepsilon)y$, we have $x = \left(1-\frac{1}{\varepsilon}\right)u + v$ and $y = \frac{u}{\varepsilon}$. Thus the Jacobian J_2 is given by

$$J_2 = \frac{\partial(x,y)}{\partial(u,v)} = \begin{pmatrix} 1-\frac{1}{\varepsilon} & 1 \\ \frac{1}{\varepsilon} & 0 \end{pmatrix} \tag{49}$$

and $\|J_2\| = \frac{1}{\varepsilon}$. Therefore, we derive the joint distribution of (U,V) as

$$\begin{aligned} f_{U,V}(u,v) &= \frac{1}{\varepsilon} f_X\left(\left(1-\frac{1}{\varepsilon}\right)u + v\right) f_Y\left(\frac{1}{\varepsilon}u\right) \\ &= \frac{1}{\varepsilon}\frac{1}{\alpha} e^{-\frac{\left(1-\frac{1}{\varepsilon}\right)u+v}{\alpha}} \frac{(u/\varepsilon)^{M-2} e^{-\frac{u}{\varepsilon\alpha}}}{\Gamma(M-1)\,\alpha^{M-1}} \\ &= \frac{2^B u^{M-2}}{\Gamma(M-1)\,\alpha^M} e^{-\frac{u+v}{\alpha}}. \end{aligned} \tag{50}$$

To guarantee that $\left(1-\frac{1}{\varepsilon}\right)u + v \ge 0$ always holds true, the condition $0 \le \frac{u}{u+v} \le \varepsilon$ is necessary.

Through comparing the results of (47) and (50), we conclude that the joint distribution of (U,V) is the same as that of (I,S).

APPENDIX B
PROOF OF LEMMA 2

The CDF of γ_b is formulated as

$$F_{\gamma_b}(t) = \Pr\{\gamma_b \le t\} = \Pr\left\{\frac{\rho_u X + \rho_u(1-\varepsilon)Y}{1+\rho_v\varepsilon Y} \le t\right\}. \tag{51}$$

To obtain the closed-form expression for $F_{\gamma_b}(t)$, we discuss the following two cases.

Case 1: When $t \ge \left(\varepsilon^{-1}-1\right)\kappa$, $\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y \ge 0$ holds true for any $y \ge 0$, and thus $\Pr\{\gamma_b \le t\}$ is derived as

$$\begin{aligned} & \Pr\{\gamma \le t\} \\ &= \int_0^\infty \Pr\left\{X \le \frac{t(1+\rho_v\varepsilon y)}{\rho_u} - (1-\varepsilon)y\right\} f_Y(y)\,dy \\ &= \int_0^\infty F_X\left(\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y\right) f_Y(y)\,dy \\ &= 1 - \int_0^\infty e^{-\frac{1}{\alpha}\left(\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y\right)} \frac{y^{M-2}e^{-y/\alpha}}{\Gamma(M-1)\,\alpha^{M-1}}\,dy \\ &= 1 - e^{-\frac{t}{\rho_u\alpha}} \int_0^\infty \frac{y^{M-2}}{\Gamma(M-1)\,\alpha^{M-1}} e^{-\frac{(\rho_v t + \rho_u)\varepsilon y}{\rho_u\alpha}}\,dy \\ &= 1 - e^{-\frac{t}{\rho_u\alpha}} \left(\frac{\rho_u}{\rho_v\varepsilon t + \rho_u\varepsilon}\right)^{M-1} \\ &\qquad \times \underbrace{\int_0^\infty \frac{y^{M-2}e^{-\frac{(\rho_v t+\rho_u)\varepsilon y}{\rho_u\alpha}}}{\Gamma(M-1)\left(\frac{\rho_u\alpha}{\rho_v\varepsilon t+\rho_u\varepsilon}\right)^{M-1}}\,dy}_{=1} \\ &= 1 - e^{-\frac{t}{\rho_u\alpha}}\beta^{M-1}, \end{aligned} \tag{52}$$

where $\beta = \frac{\rho_u}{\rho_v\varepsilon t + \rho_u\varepsilon} = \frac{\kappa\varepsilon^{-1}}{t+\kappa}$.

Case 2: When $0 \le t < \left(\varepsilon^{-1}-1\right)\kappa$, we need to restrict the condition $y \le \frac{t}{\rho_u(1-\varepsilon) - \rho_v\varepsilon t}$ to guarantee that $\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y \ge 0$. Under this case, we define $\lambda = \frac{t}{\rho_u(1-\varepsilon) - \rho_v\varepsilon t}$ and derive $\Pr\{\gamma_b \le t\}$ as

$$\begin{aligned} & \Pr\{\gamma_b \le t\} \\ &= \int_0^\lambda \Pr\left\{X \le \frac{t(1+\rho_v\varepsilon y)}{\rho_u} - (1-\varepsilon)y\right\} f_Y(y)\,dy \\ &= \int_0^\lambda F_X\left(\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y\right) f_Y(y)\,dy \\ &= F_Y(\lambda) - \int_0^\lambda e^{-\frac{1}{\alpha}\left(\frac{t}{\rho_u} + \left(\frac{\rho_v\varepsilon t}{\rho_u} + \varepsilon - 1\right)y\right)} \frac{y^{M-2}e^{-y/\alpha}}{\Gamma(M-1)\,\alpha^{M-1}}\,dy \\ &= F_Y(\lambda) - e^{-\frac{t}{\rho_u\alpha}}\beta^{M-1} \int_0^\lambda \frac{y^{M-2}e^{-\frac{y}{\alpha\beta}}}{\Gamma(M-1)\,(\alpha\beta)^{M-1}}\,dy \\ &= F_Y(\lambda) - e^{-\frac{t}{\rho_u\alpha}}\beta^{M-1} F_\Upsilon(\lambda), \end{aligned} \tag{53}$$

where Υ has a Gamma distribution $Gamma(M-1,\alpha\beta)$.

The results obtained in (52) and (53) are summarized in ***Lemma 2***.

REFERENCES

[1] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[2] P. A. Regalia, A. Khisti, Y. Liang, and S. Tomasin, "Secure communications via physical-layer and information-theoretic techniques," *IEEE Proceed.*, vol. 103, no. 10, pp. 1698–1701, Oct. 2015.

[3] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. 22, no. 6, pp. 644–654, Nov. 1976.

[4] A. D. Wyner, "The wire-tap channel," *Bell Syst. Techn. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

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

[6] A. Khisti and G. Wornell, "Secure transmission with multiple antennas–II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[7] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[8] H. Alves, R. D. Souza, M. Debbah, and M. Bennis, "Performance of transmit antenna selection physical layer security schemes," *IEEE Signal Process. Lett.*, vol. 19, no. 6, pp. 372–375, June 2012.

[9] N. Yang, P. L. Yeoh, M. Elkashlan, R. Schober, and I. B. Collings, "Transmit antenna selection for security enhancement in MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 144–154, Jan. 2013.

[10] J. Hu, Y. Cai, N. Yang, and W. Yang, "A new secure transmission scheme with outdated antenna selection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 11, pp. 2435–2446, Nov. 2015.

[11] D. Lun, H. Zhu, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no.3, pp. 1875–1888, Mar. 2010.

[12] Z. Gan, C. Li-Chia, and W. Kai-Kit, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no.3, pp. 1317–1322, Mar. 2011.

[13] H. Wang, F. Liu, and M. Yang, "Joint cooperative beamforming, jamming and power allocation to secure AF relay systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4893–4898, Oct. 2015.

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

[15] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, July 2010.

[16] X. Zhang, X. Zhou, and M. R. McKay, "On the design of artificial-noise-aided secure multi-antenna transmission in slow fading channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2170–2181, June 2013.

[17] N. Yang, S. Yan, J. Yuan, R. Malaney, R. Subramanian, and I. Land, "Artificial noise: Transmission optimization in multi-input single-output wiretap channels," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1771–1783, May 2015.

[18] T. Y. Liu, P. H. Lin, S. C. Lin, Y. W. P. Hong, and E. A. Jorswieck, "To avoid or not to avoid CSI leakage in physical layer secret communication system," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 19–25, Dec. 2015.

[19] T.-Y. Liu, S.-C. Lin, T.-H. Chang, and Y.-W. P. Hong, "How much training is enough for secrecy beamforming with artificial noise," in *Proc. IEEE ICC*, Ottawa, Canada, June 2012, pp. 4782–4787.

[20] B. He, and X. Zhou, "Secrecy on-off transmission design with channel estimation errors," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp.1923–1936, Dec. 2013.

[21] S. Bashar, Z. Ding, and Y. Li, "On secrecy of codebook-based transmission beamforming under receiver limited feedback," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1212–1223, Apr. 2011.

[22] S. C. Lin, T. H. Chang, Y. L. Liang, Y. W. P. Hong, and C. Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy with artificial noise: The noise leakage problem," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 901–915, Mar. 2011.

[23] X. Zhang, M. R. McKay, X. Zhou, and R. W. Heath Jr., "Artificial-noise-aided secure multi-antenna transmission with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2742–2754, May 2015.

[24] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[25] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[26] H. Wang, C. Wang, and W. Ng, "Artificial noise assisted secure transmission under training and feedback," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6285–6298, Dec. 2015.

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

[28] X. Zhou, M. R. McKay, B. Maham, and A. Hjørungnes, "Rethinking the secrecy outage formulation: A secure transmission design perspective," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 302–304, Mar. 2011.

[29] W. Spantipach and M. L. Honig, "Optimization of training and feedback overhead for beamforming over block fading channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6103–6115, Dec. 2010.

[30] K. 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, no. 10, pp. 2562–2579, Oct. 2003.

[31] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[32] S. Liu, Y. Hong, and E. Viterbo, "Artificial noise revisited," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3901–3911, July 2015.

[33] 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.

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

[35] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. M. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[36] J. Hu, W. Yang, N. Yang, X. Zhou, and Y. Cai, "On-off-based secure transmission design with outdated channel state information," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6075–6088, Aug. 2016.

[37] S. G. Chen, "Efficiency improvement in explicit enumeration for integer programming problems," in *Proc. IEEE IEEM*, Bangkok, Thailand, Dec. 2013, pp. 98–100.