

Physical-Layer Security in the Finite Blocklength Regime over Fading Channels

Tong-Xing Zheng, *Member, IEEE*, Hui-Ming Wang, *Senior Member, IEEE*,
Derrick Wing Kwan Ng, *Senior Member, IEEE*, and Jinhong Yuan, *Fellow, IEEE*

Abstract—This paper studies physical-layer secure transmissions from a transmitter to a legitimate receiver against an eavesdropper over slow fading channels, taking into account the impact of finite blocklength secrecy coding. A comprehensive analysis and optimization framework is established to investigate secrecy throughput for both single- and multi-antenna transmitter scenarios. Both adaptive and non-adaptive design schemes are devised, in which the secrecy throughput is maximized by exploiting the instantaneous and statistical channel state information of the legitimate receiver, respectively. Specifically, optimal transmission policy, blocklength, and code rates are jointly designed to maximize the secrecy throughput. Additionally, null-space artificial noise is employed to improve the secrecy throughput for the multi-antenna setup with the optimal power allocation derived. Various important insights are developed. In particular, 1) increasing blocklength benefits both reliability and secrecy under the proposed transmission policy; 2) secrecy throughput monotonically increases with blocklength; 3) secrecy throughput initially increases but then decreases as secrecy rate increases, and the optimal secrecy rate maximizing the secrecy throughput should be carefully chosen in order to strike a good balance between rate and decoding correctness. Numerical results are eventually presented to verify theoretical findings.

Index Terms—Physical-layer security, wiretap code, secrecy throughput, finite blocklength, optimization.

I. INTRODUCTION

In the past decade, pursuing communication security at the physical layer has received a considerable interest, e.g., [1]-[7]. In particular, physical-layer security exploits the inherent randomness of noise and wireless channels to protect wireless secure transmissions [8]-[12], which can provide an additional mechanism for security guarantee and can coexist with those security techniques already employed at the upper layers, such as key-based encipherment. Most recent progress in developing physical-layer security is motivated by Wyner's pioneering work. Specifically, the concept of secrecy capacity was first established which is defined as the supremum of

secrecy rates at which both reliability and secrecy are achieved over a wiretap channel [13]. Wyner showed that the error probability and information leakage can be made arbitrarily low concurrently with an appropriate secrecy coding, provided that a data rate below the secrecy capacity is chosen and meanwhile the data is mapped to asymptotically long codewords, i.e., the coding blocklength tends to infinity. However, the upcoming 5G wireless communication systems are required to support various novel traffic types adopting short packets to reduce the end-to-end communication latency, e.g., smart-traffic safety and machine-to-machine communications [14], [15]. For the short-packet applications, conventional physical-layer security schemes originated from infinite blocklength are generally suboptimal and the impact of finite blocklength could be destructive for secure communications. Therefore, it is necessary to rethink the analysis and design of physical-layer security for the finite blocklength regime.

A. Previous Works and Motivations

Decoding with finite blocklength will inevitably reduce the secrecy capacity and some preliminary works have been devoted to analyzing the impact of finite blocklength on secrecy for the wiretap channel. For example, the authors in [16] derived an upper bound for the information leakage probability for a given target decoding error probability demonstrating the inherent trade-off between secrecy and reliability. The authors in [17] provided both upper and lower bounds for the maximal secrecy rate capturing the impact of finite blocklength, error probability, and information leakage in both degraded discrete-memoryless wiretap channels and Gaussian wiretap channels. The obtained bounds were shown to be tighter than existing ones from [18], [19]. The work in [17] was further extended by [20], in which the optimal second-order secrecy rate was derived for a semi-deterministic wiretap channel, and the optimal tradeoff between secrecy and reliability with finite blocklength was analytically characterized. It should be noted that, all the above works were aimed to uncover the fundamental limits of secrecy performance from the information theory point of view, whereas the design of practical signaling and transmission schemes were not investigated.

In practice, due to finite blocklength penalty for practical coding schemes, even a secrecy rate below the secrecy capacity cannot guarantee a perfectly successful and secure communication. In this sense, in addition to exploring and/or improving the fundamental limits of the maximal secrecy rate, optimizing secrecy throughput seems more important from the

T.-X. Zheng is with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China, also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China, and also with the Ministry of Education Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zhengtx@mail.xjtu.edu.cn).

H.-M. Wang is with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China, and also with the Ministry of Education Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: xjbswhm@gmail.com).

D. W. K. Ng and J. Yuan are with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: w.k.ng@unsw.edu.au; j.yuan@unsw.edu.au).

perspective of transmission efficiency, particularly for fading channels where the code rates can be adapted to the fading status. Herein, the secrecy throughput denotes the amount of successfully delivered secret information subject to certain reliability and secrecy constraints. In fact, the secrecy throughput has extensively been taken as an optimization objective for the design of secure transmissions in slow fading channels in the context of infinite blocklength [21]-[25]. Nevertheless, to optimize the secrecy throughput under the constraint of finite blocklength is difficult, and the results derived for infinite blocklength, e.g., [21]-[25], cannot be directly applied. Indeed, the blocklength itself is an optimization variable, and it couples with other variables in a sophisticated manner which makes the optimization problem intractable. For instance, the authors in a recent work [26] investigated the secrecy throughput of a relay-aided secure transmission with finite blocklength, where neither the instantaneous channel state information (CSI) with respect to (w.r.t.) the legitimate receiver nor the eavesdropper is available at the transmitter side. Numerical results were presented therein to show that there exists a critical value of the blocklength that maximizes the secrecy throughput.

Despite the above endeavors, there are some fundamental questions regarding the design of physical-layer security schemes with finite blocklength that have not been thoroughly addressed. First of all, a theoretical proof of the optimal blocklength and the corresponding secrecy rate for maximizing the secrecy throughput is of great significance for the practical design of secure transmissions, which however has not yet been reported by existing literature. Also, in many applications, the transmitter is capable to acquire the instantaneous CSI of the legitimate receiver in slow fading channels via training or feedback. Yet, the potential of exploiting the instantaneous CSI to alleviate the negative impact of finite blocklength on the performance of secure communications has not been exploited. Furthermore, only the single-antenna transmitter scenario has been considered, e.g., [16]-[20], [26], and the design of the optimal signaling and code rates for multi-antenna systems with finite blocklength is still an open issue. This research work aims to provide an analytical framework and design schemes to address the abovementioned problems.

B. Contributions

This paper investigates the security issue between a pair of legitimate communicating parties in the presence of an eavesdropper, considering the impact of finite blocklength in secrecy coding. The secrecy throughput is thoroughly analyzed and optimized for both single- and multi-antenna transmitter scenarios. In particular, both adaptive and non-adaptive parameter design schemes are proposed for each scenario. The main contributions of this work are summarized as follows:

- For the single-antenna transmitter scenario, the secrecy throughput is maximized by jointly optimizing the transmission policy, blocklength, as well as code rates. Closed-form bounds and approximations for the secrecy rate are provided to facilitate the practical design of code rates for achieving a close-to-optimal performance.
- For the multi-antenna transmitter configuration, the optimality of the null-space artificial noise (AN) scheme

in terms of secrecy throughput maximization is first investigated. Afterwards, the optimal transmission policy, blocklength, code rates, and power allocation between the information-bearing signal and the AN are derived. Particularly, the power allocation and the secrecy rate are designed via the alternating optimization method, and their impacts on the system performance are further revealed.

- Numerous useful insights into the design of secure transmissions are provided with finite blocklength. For example, 1) increasing the blocklength can improve both reliability and secrecy, with properly exploiting the instantaneous CSI of the main channel and the statistical CSI of the wiretap channel, which has not been revealed by existing literature, e.g., [16]-[20]; 2) using the maximal blocklength is profitable for boosting the secrecy throughput, which is distinguished from the observation in [26]; 3) due to the finite blocklength penalty, there is a critical secrecy rate that can maximize the secrecy throughput even for the adaptive scheme, rather than always employing the maximal available secrecy rate, which is fundamentally different from the phenomenon with infinite blocklength, e.g., [21], [22].

C. Organization and Notations

The remainder of this paper is organized as below. Section II describes the system model and the underlying optimization problem. Sections III and IV detail the secrecy throughput maximization for both single- and multi-antenna transmitter scenarios. Section V draws a conclusion.

Notations: Bold lowercase letters denote column vectors. $|\cdot|$, $\|\cdot\|$, $(\cdot)^\dagger$, $(\cdot)^T$, $\ln(\cdot)$, $\mathbb{P}\{\cdot\}$, $\mathbb{E}_v[\cdot]$ denote the absolute value, Euclidean norm, conjugate, transpose, natural logarithm, probability, and the expectation over a random variable v , respectively. $f_v(\cdot)$ and $\mathcal{F}_v(\cdot)$ denote the probability density function (PDF) and cumulative distribution function (CDF) of v , respectively. $F^{-1}(\cdot)$ denotes the inverse function of a function $F(\cdot)$. $\mathcal{CN}(\mu, \sigma^2)$ denotes the circularly symmetric complex Gaussian distribution with mean μ and variance σ^2 .

II. SYSTEM MODEL AND PROBLEM DESCRIPTION

A. Channel Model

Consider a secure transmission from a transmitter (Alice) to a legitimate receiver (Bob) coexisting with an eavesdropper (Eve), as depicted in Fig. 1. Alice is equipped with $M \geq 1$ transmit antennas, whereas Bob and Eve are single-antenna devices. Quasi-static Rayleigh fading channels are considered, where the channel coherence time is on the order of the blocklength. More specifically, the fading coefficients are assumed to remain constant during the transmission of an entire codeword, but change independently and randomly between two codewords [1]. Denote the coefficients of the main and wiretap channels by \mathbf{h}_b and \mathbf{h}_e , and each entry of \mathbf{h}_b and \mathbf{h}_e follow the Gaussian distribution $\mathcal{CN}(0, \sigma_b^2)$ and $\mathcal{CN}(0, \sigma_e^2)$, respectively.¹ A common hypothesis is adopted

¹The subscripts b and e are used to refer to Bob and Eve, respectively.

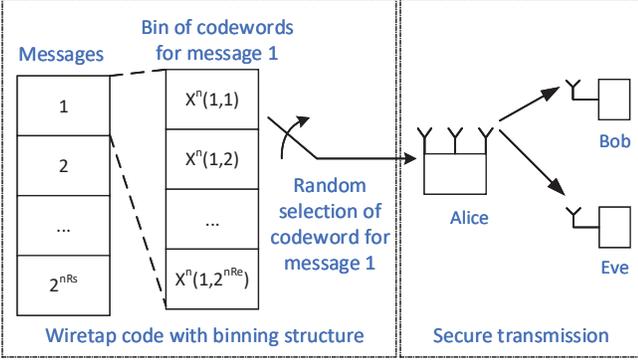


Fig. 1: Secure transmission from Alice (multi-antenna) to Bob (single-antenna) overheard by Eve (single-antenna). Alice adopts a wiretap code with a binning structure, where 2^{nR_s} messages each are mapped to a bin of 2^{nR_e} codewords with a finite blocklength n . A codeword among a set of codewords representing the same message is randomly chosen for transmission [13].

[21], [22], i.e., Bob and Eve know perfectly the instantaneous CSI of their individual channels \mathbf{h}_b and \mathbf{h}_e , and Alice has the instantaneous CSI of Bob's channel \mathbf{h}_b but does not have the instantaneous CSI of Eve's channel \mathbf{h}_e . Besides, the statistics of both channels \mathbf{h}_b and \mathbf{h}_e are available at Alice. Assume that \mathbf{h}_b , \mathbf{h}_e , and the receiver noise are mutually independent, where noise variances at Bob and Eve are denoted by w_b^2 and w_e^2 , respectively. Alice adopts a constant transmit power P . For notational simplicity, define $P_b \triangleq \frac{P}{w_b^2}$ and $P_e \triangleq \frac{P}{w_e^2}$ as the normalized power for Bob and Eve, respectively.

B. Finite Blocklength Secrecy Coding

To safeguard information confidentiality, secrecy coding should be employed to encode the secret information bits. Instead of investigating any explicit practical constructions of secrecy codes, the Wyner's wiretap code [13], as a generic code structure, is employed in this paper. A synopsis of the state-of-the-art coding schemes for wiretap channels can be found in [27].

It is reported in [1] that the Wyner's wiretap code possesses a binning structure, as illustrated in Fig. 1, where 2^{nR_s} messages are encoded to 2^{nR_t} codewords, and each message is mapped to a bin of 2^{nR_e} codewords. Here, n denotes the blocklength (i.e., the codeword length or the number of channel uses), R_s and R_t (bits/s/Hz/channel) denote the secrecy rate and codeword rate, respectively. The binning codeword rate, i.e., the rate redundancy $R_e = R_t - R_s$, reflects the cost of providing secrecy.

It is well-known that, for an infinite blocklength with $n \rightarrow \infty$, as long as the codeword rate R_t is not larger than Bob's channel capacity, Bob can recover messages with an arbitrarily low decoding error probability. On the other hand, perfect secrecy cannot always be guaranteed due to the absence of Eve's instantaneous CSI: once the rate redundancy R_e falls below Eve's channel capacity, perfect secrecy is compromised, and a secrecy outage event is said to have occurred. Nevertheless, in the finite blocklength regime which is restricted to a finite number of channel uses, no practical protocols can achieve perfectly reliable communications [28]. Hence, to capture the impact of finite blocklength, the maximal

channel coding rate for sustaining a desired decoding error probability ϵ at a finite blocklength n (e.g., $n \geq 100$) for a given signal-to-noise ratio (SNR) γ was studied in [29] and can be approximated by

$$R(\gamma, n, \epsilon) \approx C(\gamma) - \sqrt{\frac{V(\gamma)}{n}} Q^{-1}(\epsilon), \quad (1)$$

where $C(\gamma) \triangleq \log_2(1 + \gamma)$ denotes the Shannon channel capacity, $V(\gamma) \triangleq (1 - (1 + \gamma)^{-2}) \log_2^2 e$ denotes the channel dispersion [29], and $Q(x)$ is the Q -function defined as $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$. Equivalently, the decoding error probability for a given coding rate R can be expressed as

$$\epsilon(\gamma, n, R) = Q\left(\frac{C(\gamma) - R}{\sqrt{V(\gamma)/n}}\right). \quad (2)$$

For ease of notation, let $C_i \triangleq C(\gamma_i)$ and $V_i \triangleq V(\gamma_i)$ for $i \in \{b, e\}$, where γ_i denotes the corresponding SNR. Define the successful decoding probability of Bob as the complement of its decoding error probability with the codeword rate R_t . Then, the successful decoding probability conditioned on the power gain of the main channel, i.e., $\eta \triangleq \|\mathbf{h}_b\|^2$, can be expressed as

$$p_s(\eta) \triangleq 1 - \epsilon(\gamma_b, n, R_t) = 1 - Q\left(\frac{C_b - R_t}{\sqrt{V_b/n}}\right). \quad (3)$$

The secrecy performance is characterized by the information leakage probability defined below:

$$\mathcal{O}_e \triangleq \mathbb{E}_{\gamma_e} [1 - \epsilon(\gamma_e, n, R_e)]. \quad (4)$$

Remark 1: Due to the finite blocklength, the secrecy metric information leakage probability in (4) appears to be distinguished from the widely used secrecy outage probability, defined as $\mathbb{P}\{R_e \leq C_e\}$ [22], for the infinite blocklength regime with $n \rightarrow \infty$.

C. Optimization Problem

Since Alice knows Bob's instantaneous CSI perfectly, she is able to adapt the code rates to the instantaneous channel gain η , which implies that the code rates can be functions of η . This paper focuses on the metric named secrecy throughput (bits/s/Hz/channel), which measures the average successfully transmitted information bits per second per Hertz per channel use subject to a secrecy constraint $\mathcal{O}_e \leq \delta$, where $\delta \in [0, 1]$ is a pre-established threshold for the information leakage probability. Formally, the secrecy throughput is defined as

$$\mathcal{T} \triangleq \mathbb{E}_\eta [R_s(\eta)p_s(\eta)] \quad \text{s.t.} \quad \mathcal{O}_e \leq \delta, \quad (5)$$

which is averaged over η . Note that the introduction of finite blocklength leads to a different definition of secrecy throughput compared to the case of infinite blocklength which is $\mathcal{T} \triangleq \mathbb{E}_\eta [R_s(\eta)]$ [21], [22]. In addition, as will be shown later, in order to meet certain secrecy and reliability requirements during the transmission period, an on-off transmission policy

is required;² i.e., the transmission should take place only when the channel gain η exceeds some threshold $\mu > 0$. With the on-off policy, $R_s(\eta)$ is set to zero for $\eta < \mu$.

This paper aims to maximize the secrecy throughput by designing the optimal on-off threshold, signaling, blocklength, as well as code rates. The following two sections will detail the optimization for single- and multi-antenna transmitter scenarios, respectively. For each scenario, both adaptive and non-adaptive design schemes are examined, where Alice adjusts the arguments based on the instantaneous and statistical CSI of the main channel, respectively.

III. SINGLE-ANTENNA TRANSMITTER SCENARIO

For the single-antenna transmitter scenario, the SNRs of Bob and Eve are given by $\gamma_b = P_b\eta$ with $\eta = |h_b|^2$ and $\gamma_e = P_e|h_e|^2$, respectively. Clearly, γ_i is exponentially distributed with mean $\Gamma_i = P_i\sigma_i^2$ for $i \in \{b, e\}$. The subsequent two subsections aim to maximize the secrecy throughput \mathcal{T} defined in (5) by jointly designing the on-off threshold $\mu(\eta)$, the wiretap code rates $R_s(\eta)$ and $R_e(\eta)$, and the blocklength $n(\eta)$, via adaptive and non-adaptive ways, respectively. For notational convenience, these parameters are treated as functions of η by default for the adaptive scheme, with the notation η being dropped, and \mathcal{T}_A and \mathcal{T}_N are used to differentiate the adaptive scheme to its non-adaptive counterpart. The optimization problem then can be formulated as below:

$$\max_{\mu > 0, R_e > 0, R_s > 0, n} \mathcal{T} = \mathbb{E}_\eta [R_s p_s] \quad (6a)$$

$$\text{s.t. } C_b \geq R_t = R_s + R_e, \quad \forall \eta > \mu, \quad (6b)$$

$$\mathcal{O}_e \leq \delta, \quad (6c)$$

$$1 \leq n \leq N, \quad n, N \in \mathbb{Z}^+. \quad (6d)$$

Note that (6b) is interpreted as a reliability requirement since otherwise the successful decoding probability p_s in (3) falls below 0.5 and it is no better than random guessing, which is definitely not acceptable; (6c) describes the secrecy constraint; (6d) is related to a latency constraint, where the integer N denotes the maximal available blocklength imposed by a maximal tolerable delay.

A. Adaptive Optimization Scheme

In the adaptive scheme, the parameters μ , R_s , R_e , and n are designed based on η , i.e., they are adjusted in real time. A detailed optimization procedure is provided as follows.

1) *Solving R_e* : Since Q -function $Q(x)$ is a monotonically decreasing function of x , it is known that p_s defined in (3) decreases with R_e for a fixed R_s . This suggests that, the optimal R_e maximizing \mathcal{T}_A should be the minimal R_e that satisfies the secrecy constraint $\mathcal{O}_e \leq \delta$. Now that \mathcal{O}_e in (4) decreases with R_e , the optimal R_e is given as the inverse of \mathcal{O}_e at δ , i.e.:

$$R_e^* = \mathcal{O}_e^{-1}(\delta). \quad (7)$$

²The on-off policy was initially proposed for ergodic-fading channels [30], where a codeword experiences many channel realizations. It was later introduced to slow fading channels and well characterized the condition for secure transmissions [21].

Obviously, R_e^* is independent of η , but monotonically decreases with δ . This is intuitive that a larger rate redundancy is required to combat the eavesdropper in order to meet a more rigorous secrecy constraint. Although it is difficult to derive a closed-form expression for R_e^* due to the complicated Q -function, the value of R_e^* can be efficiently acquired via a bisection method with $\mathcal{O}_e(R_e) = \delta$, requiring only the computation of $Q(x)$ or a lookup table.

2) *Solving μ* : The secrecy throughput \mathcal{T}_A given in (6a) can be calculated as

$$\mathcal{T}_A = \int_{P_b\mu}^{\infty} R_s p_s f_{\gamma_b}(\gamma) d\gamma, \quad (8)$$

where $f_{\gamma_b}(\gamma) = \frac{1}{\Gamma_b} e^{-\gamma/\Gamma_b}$ is the PDF of $\gamma_b = P_b\eta$. It appears that choosing μ as small as possible is beneficial for increasing \mathcal{T}_A , on the premise of satisfying the reliability constraint (6b). In addition, constraint (6b) suggests that $C_b > R_e^* \Rightarrow \eta = \frac{\gamma_b}{P_b} > \frac{2^{R_e^*}-1}{P_b}$ must be ensured to achieve a positive R_s . Hence, the optimal on-off threshold is given by

$$\mu^* = \frac{2^{R_e^*}-1}{P_b}. \quad (9)$$

This result indicates that the transmission condition for the adaptive scheme is determined by the secrecy constraint. Apparently, μ^* is monotonically decreasing with δ since R_e^* decreases with δ . This implies, a weaker channel is still allowed for transmission for a looser secrecy constraint.

Once μ is obtained, to maximize \mathcal{T}_A in (8) only calls for maximizing $\mathcal{T}_A(\eta) \triangleq R_s p_s$ which is conditioned on η . The subproblem is described as below:

$$\max_{R_s, n} \mathcal{T}_A(\eta) = R_s p_s \quad \text{s.t. (6d), } 0 \leq R_s \leq C_b - R_e^*. \quad (10)$$

The basic idea to tackle the above problem is first to maximize p_s over n for a fixed R_s and then to design the optimal R_s that maximizes $R_s p_s$ with the optimal n .

3) *Solving n* : For any fixed $R_t \leq C_b$, there is no doubt that p_s increases with n . However, as shown in (4), $\epsilon(\gamma_e, n, R_e)$ decreases with n for $R_e \leq C_e$ but increases with n otherwise. Then, it remains unclear how \mathcal{O}_e defined in (4), as well as R_e^* in (7), varies with n . More importantly, it is less obvious if the monotonicity of p_s w.r.t. n can still hold, since $R_t = R_s + R_e^*$ becomes independent of n . Therefore, in order to derive the optimal n^* maximizing $\mathcal{T}_A(\eta)$ in (10), the monotonicity of \mathcal{O}_e or R_e^* w.r.t. n should be first identified.

Lemma 1: \mathcal{O}_e in (4) and R_e^* in (7) decrease with n .

Proof 1: Please refer to Appendix A.

Lemma 1 shows that increasing the blocklength is beneficial for decreasing the information leakage probability such that the required rate redundancy of the wiretap code can be lowered. This result is perhaps counter-intuitive, which makes sense when one realizes that a larger blocklength will yield a larger decoding error probability for Eve if Eve's channel capacity falls below the rate redundancy. With Lemma 1, the monotonicity of $\mathcal{T}_A(\eta)$ w.r.t. n is uncovered, followed by the optimal n^* that maximizes $\mathcal{T}_A(\eta)$.

Theorem 1: $\mathcal{T}_A(\eta)$ in (10) increases with n and is maximized at $n^* = N$.

Proof 2: Please refer to Appendix B.

Theorem 1 reveals that exploiting a larger blocklength is beneficial for improving the secrecy throughput under given channel gains. This result is nontrivial in light of [26] where there exists a critical value of the blocklength, instead of the maximal one, that can achieve the maximal secrecy throughput. The main reason behind the two different results lies in that, Bob's instantaneous CSI is available here and is adequately exploited, and the codeword rate will not exceed Bob's channel capacity under the on-off policy such that using a larger blocklength can always lower the decoding error probability for Bob. Combined with Lemma 1, it can be seen that increasing the blocklength improves reliability and secrecy simultaneously, thus making the secrecy throughput higher. However, this can no longer be promised in [26] where the instantaneous CSI of the main channel is unknown, and using a larger blocklength might degrade the reliability once the codeword rate exceeds Bob's channel capacity, just as implied in Lemma 1. Revisiting (8), since μ^* in the lower limit of the integral decreases with n (see (9) where R_e^* decreases with n), it is clear that the global optimal blocklength that maximizes \mathcal{T}_A is also $n^* = N$.

4) *Solving R_s :* Substituting the derived optimal R_e^* , μ^* , and n^* into (3) yields the maximal p_s , and then the optimal R_s^* can be determined by solving the following problem:

$$\max_{R_s} \mathcal{T}_A(\eta) = R_s \left[1 - Q \left(\frac{C_b - R_s - R_e^*}{\sqrt{V_b/N}} \right) \right] \quad (11a)$$

$$\text{s.t. } 0 < R_s \leq C_b - R_e^*. \quad (11b)$$

Theorem 2: $\mathcal{T}_A(\eta)$ in (11) is a concave function of R_s , and its maximal value is achieved at

$$R_s^* = \begin{cases} C_b - R_e^*, & \eta \leq \frac{\gamma_b^\circ}{P_b}, \\ R_s^\circ, & \text{otherwise,} \end{cases} \quad (12)$$

where $\gamma_b^\circ \in \left(\sqrt{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\pi}{2N}}} - 1, e^{\frac{\sqrt{\pi V_b} + R_e^* \ln 2}{2N}} - 1 \right)$ is

the unique root $\gamma_b > 0$ that satisfies $C_b - \sqrt{\frac{\pi V_b}{2N}} = R_e^*$, and R_s° is the unique zero-crossing $R_s < C_b - R_e^*$ of the derivative

$$\frac{d\mathcal{T}_A(\eta)}{dR_s} = 1 - Q \left(\frac{C_b - R_s - R_e^*}{\sqrt{V_b/N}} \right) - \frac{R_s \sqrt{N}}{\sqrt{2\pi V_b}} e^{-\frac{(C_b - R_s - R_e^*)^2}{2V_b/N}}. \quad (13)$$

Proof 3: Please refer to Appendix C.

Theorem 2 presents an optimal secrecy rate R_s^* that differs from the one for infinite blocklength with $N \rightarrow \infty$, where in the latter employing the maximal achievable secrecy rate $R_s^* = C_b - R_e^*$ is always optimal for secrecy throughput improvement. The fundamental reason behind such difference lies in the decoding failure caused by finite blocklength. Specifically, when the quality of the main channel is poor (i.e., a small η) or when a large rate redundancy R_e^* is required, e.g., due to a high average SNR of Eve or a stringent secrecy requirement, the successful decoding probability p_s is initially small and decreases slowly with R_s . In this case, the secrecy throughput improvement is mainly bottlenecked by R_s , and

hence it is necessary to choose the maximal secrecy rate $R_s^* = C_b - R_e^*$. Otherwise, p_s is initially large but drops rapidly with R_s , thus dramatically degrading the secrecy throughput. Therefore, a relatively small R_s is supposed to be chosen to strike a good balance between the decoding and throughput performance.

The optimal secrecy rate $R_s^* \leq C_b - R_e^*$ in (12) can be obtained efficiently using the Newton's method, despite its implicit form. The following corollaries further give a closed-form asymptotically tight lower bound R_s^L on R_s^* and provide useful insights into the behavior of R_s^* .

Corollary 1: The optimal secrecy rate R_s^* in (12) satisfies

$$R_s^* \geq R_s^L \triangleq C_b - R_e^* - \sqrt{\frac{2V_b}{N} \ln \left(\frac{1}{2} + \frac{C_b - R_e^*}{\sqrt{2\pi V_b/N}} \right)}. \quad (14)$$

Proof 4: The result follows by finding a lower bound on $\frac{d\mathcal{T}_A(\eta)}{dR_s}$ in (13) applying the inequalities $Q(x) \leq \frac{1}{2}e^{-x^2/2}$ and $R_s \leq C_b - R_e^*$ and then setting the resultant lower bound to zero.

The term $\sqrt{\frac{2V_b}{N} \ln \left(\frac{1}{2} + \frac{C_b - R_e^*}{\sqrt{2\pi V_b/N}} \right)}$ in (14) is interpreted as the secrecy rate loss arisen from finite blocklength. This term vanishes as $N \rightarrow \infty$ or $R_e^* \rightarrow C_b - \sqrt{\frac{\pi V_b}{2N}}$, and accordingly R_s^* approaches $C_b - R_e^*$. In this sense, the lower bound R_s^L can be employed as a computational convenient alternative to the optimal R_s^* , particularly for the large blocklength scenarios.

Corollary 2: The optimal secrecy rate R_s^* monotonically increases with the channel gain η .

Proof 5: It is proved that $\frac{C_b - R_s - R_e^*}{\sqrt{V_b}}$ in (13) increases with η such that $\frac{d\mathcal{T}_A(\eta)}{dR_s}$ increases with η . Then, using the derivative rule for implicit functions with $\frac{d\mathcal{T}_A(\eta)}{dR_s} = 0$ reaches $\frac{dR_s^*}{d\eta} > 0$.

Fig. 2 depicts secrecy throughput $\mathcal{T}_A(\eta)$ versus secrecy rate R_s for different blocklength N and channel gain η . The concavity of $\mathcal{T}_A(\eta)$ on R_s given by Theorem 2 is well verified. Specifically, $\mathcal{T}_A(\eta)$ first increases and then decreases with R_s , and there exists an optimal R_s^* that maximizes $\mathcal{T}_A(\eta)$. It is also found that $\mathcal{T}_A(\eta)$ almost linearly increases with R_s at first, since the throughput loss due to decoding error is negligible. Note that the curves in the figure are cut in different points which represent different values of the maximal achievable secrecy rate R_s^{\max} for different N and η , and it is obvious that R_s^{\max} increases with N and η . As η grows, $\mathcal{T}_A(\eta)$ improves significantly and the corresponding optimal R_s^* increases, which validates Corollary 2. The underlying reason is that, when the main channel quality improves, choosing a larger R_s contributes more to improving $\mathcal{T}_A(\eta)$ compared with increasing the successful decoding probability p_s (by lowering R_s). In addition, as proved in Theorem 1, $\mathcal{T}_A(\eta)$ increases with N . It is also proved that the optimal R_s^* increases with N as $\eta \rightarrow \infty$. However, it is no longer true when η is too small, e.g., $\eta = 3$ dB. This is because, for a low channel quality, the decoding performance becomes a key restricting factor on throughput improvement, and hence R_s should be decreased to ensure a large p_s as N increases. Moreover, the secrecy throughput obtained with the lower bound R_s^L in Corollary 1 approaches closely the optimal one particularly when N is

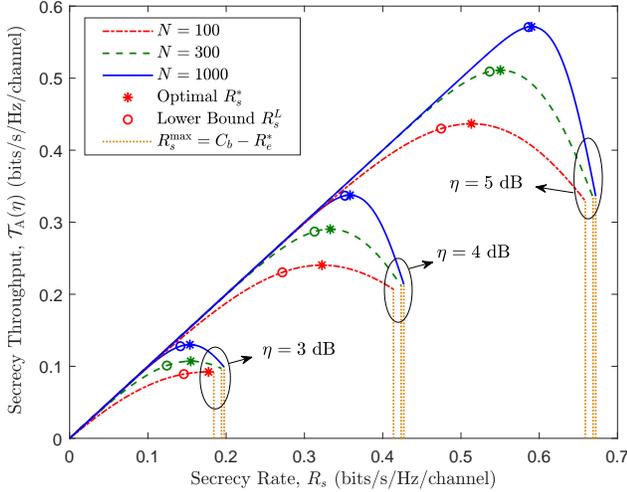


Fig. 2: $\mathcal{T}_A(\eta)$ vs. R_s for different N and η , with $P_b = 0$ dB, $\Gamma_e = 0$ dB, and $\delta = 0.2$.

sufficiently large, which demonstrates the usefulness of the lower bound.

B. Non-Adaptive Optimization Scheme

This section devises a non-adaptive optimization scheme where the parameters μ , R_s , R_e , and n are designed based on the statistical CSI of the main channel and remain unchanged during the transmission period. Such a non-adaptive scheme can be computed off-line, which significantly lowers the complexity compared with an adaptive one.

Since all the parameters are independent of the channel gain η , the problem of maximizing the secrecy throughput in (5) can be recast as follows:

$$\max_{\mu, R_e, R_s, n} \mathcal{T}_N = R_s \bar{p}_s \quad \text{s.t.} \quad (6b) - (6d), \quad (15)$$

where $\bar{p}_s = \int_{P_b \mu}^{\infty} p_s f_{\gamma_b}(\gamma) d\gamma$ denotes the average successful decoding probability.

The above problem can be handled via similar steps for its adaptive counterpart in Sec. III-A. To begin with, in order to increase p_s for a given R_s , a minimal rate redundancy R_e should be chosen while satisfying the secrecy constraint $\mathcal{O}_e \leq \delta$. Hence, the optimal R_e^* is given in (7). It can be inferred from (15) that a smaller transmission threshold μ can produce a larger \mathcal{T}_N . Nonetheless, $C_b \geq \log_2(1 + P_b \mu) \geq R_s + R_e^*$ must be ensured, since otherwise there would always exist a transmission initiated when $\eta > \mu$ while violating the reliability constraint (6b). Consequently, the optimal μ^* for a fixed R_s is given by

$$\mu^* = \frac{2^{R_s + R_e^*} - 1}{P_b}. \quad (16)$$

Note that in order to support a constant secrecy rate R_s , the optimal on-off threshold μ^* for the non-adaptive scheme is generally larger than that of the adaptive one as given in (9). On the other hand, the optimal μ^* monotonically decreases with δ and n , which is similar to the adaptive case. That is to

say, the transmission condition can be relaxed when facing a looser secrecy requirement or using a larger blocklength.

Substituting R_e^* and μ^* into \bar{p}_s and invoking the approximation of Q -function in (49) yields

$$\begin{aligned} \bar{p}_s &= \int_{P_b \mu^*}^{\infty} [1 - \Xi(\gamma_b, n, R_s + R_e^*)] f_{\gamma_b}(\gamma) d\gamma \\ &\stackrel{(a)}{=} 1 - \mathcal{F}_{\gamma_b}(\theta_b^2) \int_{\theta_b^2}^{\tau_b^u} \left(\frac{1}{2} - \frac{\beta}{\theta_b} (\gamma - \theta_b^2) \right) f_{\gamma_b}(\gamma) d\gamma \\ &\stackrel{(b)}{=} 1 - \frac{1}{2} \mathcal{F}_{\gamma_b}(\theta_b^2) - \frac{\beta}{\theta_b} \int_{\theta_b^2}^{\tau_b^u} \mathcal{F}_{\gamma_b}(\gamma) d\gamma, \end{aligned} \quad (17)$$

where (a) is due to $\theta_b = \sqrt{P_b \mu^*} = \sqrt{2^{R_s + R_e^*} - 1}$, $\beta = \frac{\sqrt{n}}{2\pi}$, and $\tau_b^u = \theta_b^2 + \frac{\theta_b}{2\beta}$, and (b) stems from the use of partial integration. With (17), the problem of maximizing \mathcal{T}_N over n and R_s can be equivalently transformed as below:

$$\max_{\beta, \theta_b} \mathcal{T}_N = [\log_2(1 + \theta_b^2) - R_e^*] \bar{p}_s \quad (18a)$$

$$\text{s.t.} \quad \frac{1}{2\pi} \leq \beta \leq \frac{\sqrt{N}}{2\pi}, \quad \theta_b > \sqrt{2^{R_e^*} - 1}. \quad (18b)$$

Theorem 3: \mathcal{T}_N in (18) is a monotonically increasing function of β or n .

Proof 6: The result follows by proving that

$$\begin{aligned} \frac{d\mathcal{T}_N}{d\beta} &= -\frac{dR_e^*}{d\beta} \bar{p}_s + [\log_2(1 + \theta_b^2) - R_e^*] \frac{d\bar{p}_s}{d\beta} \\ &\stackrel{(a)}{>} [\log_2(1 + \theta_b^2) - R_e^*] \frac{d\bar{p}_s}{d\beta} \stackrel{(b)}{>} 0, \end{aligned} \quad (19)$$

where (a) is due to $\frac{dR_e^*}{dn} < 0$ from (51), and (b) follows from $\frac{d\bar{p}_s}{d\beta} = \frac{1}{\theta_b} \int_{\theta_b^2}^{\tau_b^u} [\mathcal{F}_{\gamma_b}(\tau_b^u) - \mathcal{F}_{\gamma_b}(\gamma)] d\gamma > 0$ as $\mathcal{F}_{\gamma_b}(\gamma)$ is an increasing function of γ .

Theorem 3 suggests that Alice should use the maximal blocklength to maximize the secrecy throughput for the non-adaptive scheme, regardless of other parameters, i.e., the globally optimal blocklength is $n^* = N$. More importantly, this conclusion holds for any distribution of γ_b .

Substituting the CDF $\mathcal{F}_{\gamma_b}(\gamma) = 1 - e^{-\gamma/\Gamma_b}$ into (17) yields

$$\mathcal{T}_N = \frac{1}{2} [\log_2(1 + \theta_b^2) - R_e^*] [1 + Y(\theta_b)] e^{-\frac{\theta_b^2}{\Gamma_b}}, \quad (20)$$

where $Y(\theta_b) = \frac{2\beta\Gamma_b}{\theta_b} (1 - e^{-\frac{\theta_b}{2\beta\Gamma_b}}) > 0$. The optimal θ_b^* that maximizes \mathcal{T}_N is provided below.

Theorem 4: \mathcal{T}_N in (20) is first-increasing-then-decreasing w.r.t. θ_b ; the optimal θ_b^* maximizing \mathcal{T}_N is the unique root $\theta_b > \sqrt{2^{R_e^*} - 1}$ of $G(\theta_b) = 0$, where $G(\theta_b)$ is a decreasing function of θ_b :

$$G(\theta_b) = \frac{1 + Y(\theta_b)}{\ln 2} - [\log_2(1 + \theta_b^2) - R_e^*] \frac{1 + \theta_b^2}{\theta_b} g(\theta_b), \quad (21)$$

with $g(\theta_b) = \left(\frac{1}{2\theta_b} + \frac{1}{4\beta\Gamma_b} + \frac{\theta_b}{\Gamma_b} \right) Y(\theta_b) + \frac{\theta_b}{\Gamma_b} - \frac{1}{2\theta_b}$.

Proof 7: Please refer to Appendix D.

Based on Theorem 4, the optimal θ_b^* or secrecy rate $R_s^* = \log_2(1 + (\theta_b^*)^2) - R_e^*$ can be efficiently calculated using a bisection search with $G(\theta_b) = 0$, and thus the maximal \mathcal{T}_N^* can be obtained from (18). The following corollaries demonstrate

the behavior of R_s^* w.r.t. to the average channel power gain $\sigma_b^2 = \frac{\Gamma_b}{P_b}$ and provide a closed-form approximation of R_s^* at the large σ_b^2 regime.

Corollary 3: The optimal R_s^* monotonically increases with σ_b^2 .

Proof 8: Following similar steps as the proof of Theorem 4, it can be verified that $G(\theta_b)$ in (21) increases with σ_b^2 such that $\frac{dR_s^*}{d\sigma_b^2} = -\frac{\partial G(\theta_b)/\partial \sigma_b^2}{\partial G(\theta_b)/\partial R_s^*} > 0$, which completes the proof.

Corollary 3 suggests that a larger secrecy rate should be employed to boost the secrecy throughput when the quality of the main channel improves, despite the fact that it might deteriorate the decoding correctness at Bob.

Corollary 4: At the regime of $\sigma_b^2 \rightarrow \infty$, the optimal secrecy rate R_s^* is approximated by

$$\begin{aligned} R_s^* &\approx R_s^A = \log_2(e) \mathcal{W}_0 \left(\sigma_b^2 2^{-R_e^*} \right) \\ &\approx \log_2(\sigma_b^2) - R_e^* - \log_2 \left[\ln(\sigma_b^2) - R_e^* \ln 2 \right], \end{aligned} \quad (22)$$

where $\mathcal{W}_0(x)$ is the Lambert's W function [38, Sec. 4.13] that satisfies $x = \mathcal{W}_0(x)e^{\mathcal{W}_0(x)}$.

Proof 9: It is clear that $Y(\theta_b) \rightarrow 1$ and $g(\theta_b) \rightarrow \frac{2\theta_b}{\Gamma_b}$ as $\sigma_b^2 \rightarrow \infty$. Substituting the results into (21) with $\theta_b^2 = 2^{R_s+R_e^*} - 1$ and letting $G(\theta_b) = 0$ produce the first approximation. The second approximation comes from the expansion of $\mathcal{W}_0(x)$ as $x \rightarrow \infty$ that $\mathcal{W}_0(x) \approx \ln x - \ln(\ln x)$.

Fig. 3 plots the secrecy throughput \mathcal{T}_N versus the secrecy rate R_s for different values of the blocklength N and the average channel gain σ_b^2 . It can be seen that \mathcal{T}_N first increases and then decreases with R_s , which validates Theorem 4. The optimal R_s^* maximizing \mathcal{T}_N increases with σ_b^2 , which verifies Corollary 3 well, and the reason behind is similar to that for Corollary 2. It can also be observed that the optimal R_s^* is almost impervious to different N . This is because, the optimal secrecy rate for the non-adaptive scheme only depends on the average successful decoding probability, and the averaging process softens the impact of the blocklength. Theorem 3 is also confirmed, where it is found that \mathcal{T}_N increases with N . In addition, the secrecy throughput with the approximate R_s^A obtained in Corollary 4 is almost coincided with that of the optimal R_s^* , which demonstrates the practicability of the low-complexity approximation.

Fig. 4 compares the secrecy throughput for adaptive and non-adaptive schemes with different blocklength N . The left-hand-side figure depicts the maximal secrecy throughput \mathcal{T}^* , where \mathcal{T}_A^* for the adaptive case improves as N increases whereas \mathcal{T}_N^* for the non-adaptive case almost remains unchanged. When the average channel gain σ_b^2 increases or the secrecy constraint becomes relaxed (i.e., a larger δ), the maximal \mathcal{T}^* for both schemes improves significantly, and the gap $\mathcal{T}_A^* - \mathcal{T}_N^*$ increases. The right-hand-side figure illustrates the relative throughput gain $\Delta\mathcal{T} \triangleq \frac{\mathcal{T}_A^* - \mathcal{T}_N^*}{\mathcal{T}_N^*}$ which reflects the superiority of the adaptive scheme over its non-adaptive counterpart. It is shown that $\Delta\mathcal{T}$ grows dramatically with N but decreases with σ_b^2 and δ . This suggests that the adaptive scheme is more preferred for some *unfavorable* scenarios, e.g., with a large blocklength (large delay), a poor channel quality, or a stringent secrecy requirement; otherwise, the non-

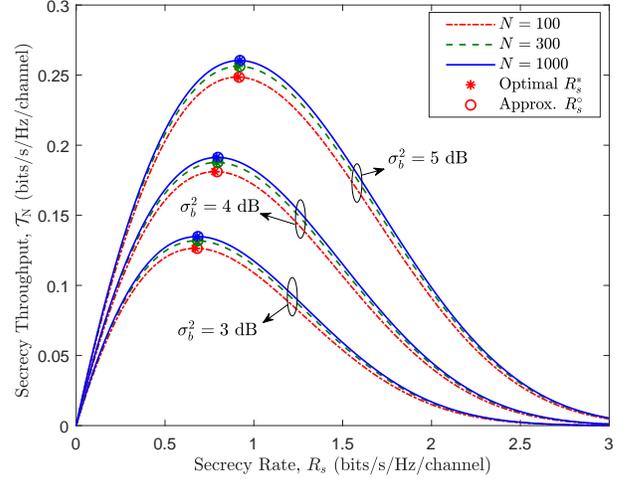


Fig. 3: \mathcal{T}_N vs. R_s for different N and σ_b^2 , with $P_b = 0$ dB, $\Gamma_e = 0$ dB, and $\delta = 0.2$.

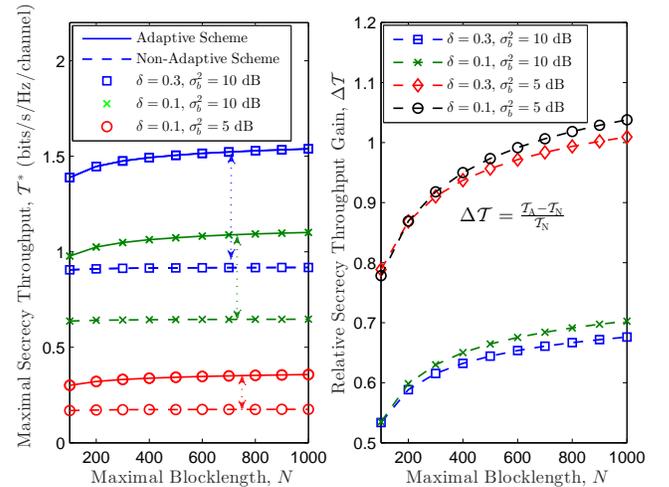


Fig. 4: \mathcal{T}^* and $\Delta\mathcal{T}$ vs. N for different σ_b^2 and δ , with $P_b = 0$ dB and $\Gamma_e = 0$ dB.

adaptive scheme could be an alternative choice owing to its low implementation complexity.

IV. MULTI-ANTENNA TRANSMITTER SCENARIO

When Alice is equipped with multiple antennas, she can intentionally transmit AN together with the information-bearing signal to degrade Eve's channel quality. Generally, the null-space AN scheme, in which the AN is injected uniformly in directions orthogonal to the main channel, is heuristically employed in the context of infinite blocklength [33]. The near-optimality of AN in terms of improving secrecy capacity for the multi-input single-output wiretap channel was first proved in [34] from a rigorous information-theoretic perspective, and its degraded performance was later observed for the multi-input multi-output wiretap channel [35]. On the other hand, it was argued in [36] that distributing a certain proportion of AN in the direction of main channel can surprisingly gain a larger ergodic secrecy rate. When it comes to finite blocklength, since

decoding failure might occur even when the codeword rate lies below the channel capacity, which is quite different from the infinite blocklength case, it is still unclear whether the null-space AN is optimal and how the optimal power allocation of the AN scheme should be determined for maximizing the secrecy throughput. To this end, this section focuses on the optimization of secrecy throughput with finite blocklength for the multi-antenna scenario, where the optimality of the null-space AN scheme will be identified first.

Considering a general scenario where the AN is not restricted to be orthogonal to the main channel, Alice's transmitted signal can be constructed in the form of

$$\mathbf{x} = \sqrt{\phi P} \mathbf{w} (\sqrt{\alpha} s + \sqrt{1-\alpha} v) + \sqrt{\frac{(1-\phi)P}{M-1}} \mathbf{W}_\perp \mathbf{z}, \quad (23)$$

where $\mathbf{w} = \frac{\mathbf{h}_b^\dagger}{\|\mathbf{h}_b\|}$ denotes the beamforming vector for the main channel, \mathbf{W}_\perp denotes the $M \times (M-1)$ projection matrix onto the null space of \mathbf{h}_b such that $\mathbf{h}_b^\top \mathbf{W}_\perp = \mathbf{0}$, and the columns of $[\mathbf{w} \ \mathbf{W}_\perp]$ constitute an orthogonal basis; s , v , and \mathbf{z} denote the information signal, the AN in the direction of \mathbf{w} , and the AN in the null space \mathbf{W}_\perp , with each element obeying $\mathcal{CN}(0, 1)$; $\phi \in [0, 1]$ represents the fraction of the total transmit power P allocated to the direction of \mathbf{w} , and $\alpha \in [0, 1]$ represents the power allocation ratio of the information signal to ϕP . With (23), the received signal-to-interference-plus-noise ratios (SINRs) at Bob and Eve are respectively

$$\gamma_b = \frac{\alpha \phi P_b \eta}{(1-\alpha) \phi P_b \eta + 1}, \quad (24)$$

$$\gamma_e = \frac{\alpha \phi P_e \|\mathbf{h}_e^\top \mathbf{w}\|^2}{(1-\alpha) \phi P_e \|\mathbf{h}_e^\top \mathbf{w}\|^2 + \frac{(1-\phi)P_e \|\mathbf{h}_e^\top \mathbf{W}_\perp\|^2}{M-1} + 1}, \quad (25)$$

where $\eta = \|\mathbf{h}_b\|^2$. The successful decoding probability p_s and the information leakage probability \mathcal{O}_e for the multi-antenna case are still given by (2) and (4), respectively. The corresponding secrecy throughput optimization problem can be formulated as below:

$$\max_{\mu, R_e, R_s, n, \alpha, \phi} \mathcal{T} = \mathbb{E}_\eta [R_s p_s] \quad \text{s.t.} \quad (6b) - (6d), \quad 0 \leq \alpha, \phi \leq 1. \quad (26)$$

The following subsections will first detail the optimization procedure for both adaptive and non-adaptive schemes, and then briefly discuss the scenario of a multi-antenna Eve.

A. Adaptive Optimization Scheme

This subsection optimizes the secrecy throughput \mathcal{T}_A by designing the parameters involved in problem (26) adaptively according to the instantaneous channel realization \mathbf{h}_b .

1) *Solving R_e* : Similar to the single-antenna case, the optimal rate redundancy is given by $R_e^* = \mathcal{O}_e^{-1}(\delta)$ with \mathcal{O}_e in (4). Note that R_e^* herein is a function of ϕ and α .

2) *Solving α* : Resort to a function $\kappa(x, \alpha) \triangleq \frac{x\alpha}{x(1-\alpha)+1}$ defined in [37], which increases with x for $\alpha > 0$. Then, the SINRs γ_b in (24) and γ_e in (25) can be reformulated as $\gamma_b(\phi, \alpha) = \kappa(\gamma_b(\phi, 1), \alpha)$ and $\gamma_e(\phi, \alpha) = \kappa(\gamma_e(\phi, 1), \alpha)$. Define $\Phi_e(\phi, \alpha) \triangleq 2^{R_e^*} - 1$ as the SINR threshold for $\gamma_e(\phi, \alpha)$ such that $R_e^* = \log_2(1 + \Phi_e(\phi, \alpha))$. Recalling the secrecy

constraint $\mathcal{O}_e(\Phi_e; \theta, \alpha) = \delta$, $\Phi_e(\phi, \alpha)$ is the δ -upper quantile of $\gamma_e(\phi, \alpha)$ such that it also follows the form $\Phi_e(\phi, \alpha) = \kappa(\Phi_e(\phi, 1), \alpha)$ [37]. Hence, the condition for guaranteeing a positive secrecy rate is described as

$$\begin{aligned} \gamma_b(\phi, \alpha) > \Phi_e(\phi, \alpha) &\Rightarrow \kappa(\gamma_b(\phi, 1), \alpha) > \kappa(\Phi_e(\phi, 1), \alpha) \\ &\Rightarrow \gamma_b(\phi, 1) > \Phi_e(\phi, 1) \\ &\stackrel{(a)}{\Rightarrow} \rho_b > \rho_e(\phi), \end{aligned} \quad (27)$$

where $\rho_b \triangleq P_b \eta$, $\rho_e(\phi) \triangleq \frac{\Phi_e(\phi, 1)}{\phi}$, and (a) is due to $\gamma_b(\phi, 1) = \phi P \eta$. Then, the threshold μ can be simply set as $\mu(\phi) = \frac{\rho_e(\phi)}{P_b}$ for any fixed ϕ . Revisiting (8), since $\mu(\phi)$ is independent of α , the optimal α^* that maximizes \mathcal{T}_A can be obtained by maximizing $\mathcal{T}_A(\eta) = R_s p_s$, where p_s is defined in (3) and can be rewritten as

$$p_s = 1 - Q\left(\sqrt{n} \lambda_b \frac{\ln \lambda_b - \ln \lambda_e - R_s \ln 2}{\sqrt{\lambda_b^2 - 1}}\right), \quad (28)$$

with $\lambda_b \triangleq 1 + \kappa(\gamma_b(\phi, 1), \alpha) > \lambda_e \triangleq 1 + \kappa(\Phi_e(\phi, 1), \alpha) > 1$. Although it is difficult to see how p_s varies with α for a fixed $R_s < \log_2 \frac{\lambda_b}{\lambda_e}$ as both λ_b and λ_e increase with α , the following theorem provides the optimal α^* that maximizes \mathcal{T}_A .

Theorem 5: $\alpha^* = 1$ is optimal for maximizing the secrecy throughput \mathcal{T}_A .

Proof 10: Please refer to Appendix E.

Theorem 5 suggests that there is no need to inject the AN in the main channel direction for secrecy throughput improvement with finite blocklength. The reason is that, once the main channel quality suffices to guarantee $\lambda_b > \lambda_e$, a larger α can improve the term $\frac{\ln \lambda_b - \ln \lambda_e}{\sqrt{\lambda_b^2 - 1}}$ in (28) which reflects the channel superiority of the main channel over the wiretap channel.

Define $\xi \triangleq \frac{\phi^{-1}-1}{M-1}$. Substituting $\alpha^* = 1$ into (24) and (25) yields the CDFs of γ_b and γ_e :

$$\mathcal{F}_{\gamma_b}(\gamma) = 1 - e^{-\frac{\gamma}{\phi \Gamma_b}} \sum_{k=0}^{M-1} \frac{1}{k!} \left(\frac{\gamma}{\phi \Gamma_b}\right)^k, \quad (29)$$

$$\mathcal{F}_{\gamma_e}(\gamma) = 1 - e^{-\frac{\gamma}{\phi \Gamma_e}} (1 + \xi \gamma)^{1-M}, \quad (30)$$

3) *Solving μ* : The threshold $\mu(\phi) = \frac{\rho_e(\phi)}{P_b}$ mentioned in the last step is related to ϕ . This step further determines the optimal μ^* which is independent of ϕ and η . For tractability, consider an asymptotically large blocklength and exploit the tail property of the Q -function, then the information leakage probability \mathcal{O}_e is approximated as [32]

$$\mathcal{O}_e(\Phi_e) \approx e^{-\frac{\Phi_e}{\phi \Gamma_e}} (1 + \xi \Phi_e)^{1-M}. \quad (31)$$

Fig. 5 shows that the approximate $\mathcal{O}_e(\Phi_e)$ is extremely close to the exact value for quite a wide range of ϕ , M , n , and Γ_e , and it then can be adopted to facilitate the subsequent analysis and optimization. Revisiting $\rho_e(\phi) = \frac{\Phi_e(\phi)}{\phi}$ with $\mathcal{O}_e(\Phi_e(\phi)) = \delta$, the following lemma is obtained.

Lemma 2 ([37]): $\rho_e(\phi) > 0$, $\frac{d\rho_e(\phi)}{d\phi} = \frac{\rho_e(\phi)}{[1+\phi\rho_e(\phi)\xi]/\Gamma_e+1-\phi} > 0$, and $\frac{d^2\rho_e(\phi)}{d\phi^2} > \frac{2}{\rho_e(\phi)} \left[\frac{d\rho_e(\phi)}{d\phi}\right]^2 > 0$.

Lemma 2 indicates that $\rho_e(\phi)$ increases with ϕ . It is observed from (27) that no positive R_s can be achieved if

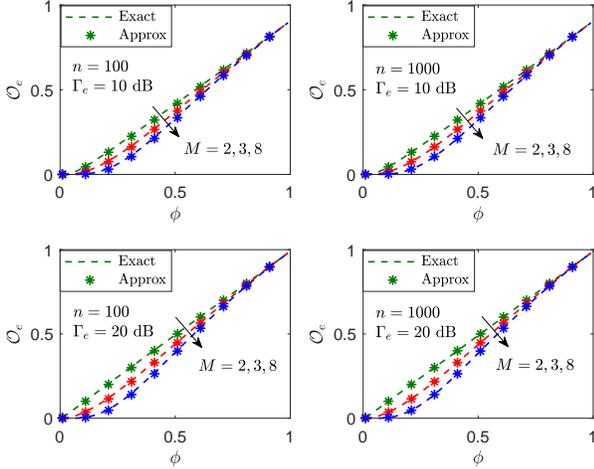


Fig. 5: \mathcal{O}_e vs. ϕ for different M , n , and Γ_e .

$P_b\eta \leq \rho_e(0)$. To avoid this, the optimal on-off threshold should be chosen as

$$\mu^* = \frac{\rho_e(0)}{P_b}. \quad (32)$$

4) *Solving n* : This step gives the optimal blocklength n^* that maximizes secrecy throughput.

Theorem 6: $n^* = N$ is optimal for maximizing \mathcal{T}_A in (26).

Proof 11: The proof is similar to that of Theorem 1. According to (53), one only needs to prove that $f_{\gamma_e}(\tau_e^l) > f_{\gamma_e}(\tau_e^u)$. The PDF of γ_e is calculated from (30) and is given by

$$f_{\gamma_e}(\gamma) = \left(\frac{1}{\phi\Gamma_e(1+\xi\gamma)^{M-1}} + \frac{\xi(M-1)}{(1+\xi\gamma)^M} \right) e^{-\frac{\gamma}{\phi\Gamma_e}}. \quad (33)$$

Apparently, $f_{\gamma_e}(\gamma)$ decreases with γ such that $f_{\gamma_e}(\tau_e^l) > f_{\gamma_e}(\tau_e^u)$, which completes the proof.

Theorem 6 suggests that a multi-antenna transmitter also should adopt the maximal blocklength to maximize the secrecy throughput, regardless of the power allocation and code rates. This is validated by Fig. 6, and the reason behind is similar to that of the single-antenna case.

5) *Solving ϕ* : By now, the secrecy throughput $\mathcal{T}_A(\eta)$ conditioned on η is given by

$$\mathcal{T}_A(\eta) = R_s \left(1 - Q \left[\sqrt{N}\lambda_b \frac{\ln \frac{\lambda_b}{\lambda_e} - R_s \ln 2}{\sqrt{\lambda_b^2 - 1}} \right] \right), \quad (34)$$

where $\lambda_b = 1 + \phi\rho_b$ and $\lambda_e = 1 + \phi\rho_e$ with $\rho_b > \rho_e$. For notational simplicity, ϕ has been dropped from $\rho_e(\phi)$. Obviously, maximizing $\mathcal{T}_A(\eta)$ is equivalent to maximizing the following function:

$$L(\phi) = \frac{\lambda_b}{\sqrt{\lambda_b^2 - 1}} \left(\ln \frac{\lambda_b}{\lambda_e} - R_s \ln 2 \right). \quad (35)$$

Theorem 7: $L(\phi)$ in (35) is a concave function of ϕ , and the optimal ϕ^* maximizing $L(\phi)$ is

$$\phi^* = \begin{cases} 1, & \eta \geq \frac{\rho_e^0}{P_b} \text{ and } \frac{\rho_e(1)}{1+\rho_e(1)} < \frac{1}{1+\Gamma_e}, \\ \phi^\circ, & \text{otherwise.} \end{cases} \quad (36)$$

Here ϕ° is the unique zero-crossing $\phi \in [0, 1)$ of the following derivative:

$$\frac{dL(\phi)}{d\phi} = \frac{(1 - A_\phi)\lambda_b - 1}{\phi\sqrt{\lambda_b^2 - 1}} - \frac{(\lambda_b - 1)(\ln \lambda_b - B_\phi)}{\phi(\lambda_b^2 - 1)^{3/2}}, \quad (37)$$

where $A_\phi \triangleq \frac{\phi}{\lambda_e} \left(\rho_e + \phi \frac{d\rho_e}{d\phi} \right)$ and $B_\phi \triangleq \ln \lambda_e + R_s \ln 2$ with $\frac{d\rho_e}{d\phi}$ given in Lemma 2, and ρ_b° is the unique root ρ_b of the equation $X(\rho_b) = 0$ with $X(\rho_b)$ given below:

$$X(\rho_b) = (1 - A_1)(1 + \rho_b) - 1 - \frac{\ln(1 + \rho_b) - B_1}{2 + \rho_b}. \quad (38)$$

Proof 12: Please refer to Appendix F.

Theorem 7 shows that, the naive beamforming scheme without injecting any AN is optimal for maximizing the secrecy throughput only when the quality of the main channel is good enough and meanwhile the quality of the wiretap channel is poor or a high information leakage probability is acceptable. Using the derivative rule for implicit functions with (37) proves that $\frac{d\phi^*}{dR_s} > 0$, which suggests that in order to support a higher secrecy rate, a larger fraction of power should be allocated to the information signal although at the cost of a larger required rate redundancy.

For a robust design perspective, a worst-case scenario is considered by ignoring Eve's thermal noise, i.e., $\Gamma_e \rightarrow \infty$ in (31), such that $\rho_e = \frac{\Lambda}{1-\phi}$ with $\Lambda = (M-1)(\delta^{\frac{1}{1-M}} - 1)$. It is seen from (37) that ϕ^* is a function of η and δ , and the monotonicity of ϕ^* is revealed as below.

Corollary 5: For the worst case $\Gamma_e \rightarrow \infty$, the optimal power allocation ϕ^* is non-decreasing w.r.t. η and δ . Moreover, $\lim_{\eta \rightarrow \infty} \phi^* = \frac{1}{\sqrt{\Lambda+1}}$ and $\lim_{\delta \rightarrow 1} \phi^* = 1$.

Proof 13: Please refer to Appendix G.

Corollary 5 suggests that when the quality of the main channel improves (i.e., a larger η) or the secrecy requirement is relaxed (i.e., a larger δ), it would be more appealing to use a higher signal power to promote the main channel than to increase the AN power to degrade the wiretap channel. This is because that the main channel becomes the dominate factor to the improvement of secrecy throughput. Different from Theorem 7 where $\phi^* = 1$ can be achieved, the optimal ϕ^* here only can be increased up to $\frac{1}{\sqrt{\Lambda+1}}$ as $\eta \rightarrow \infty$ due to Eve's background noise being ignored. Besides, it is unsurprising that $\phi^* = 1$ for $\delta = 1$ since there is no secrecy requirement.

6) *Solving R_s* : For any given power allocation ϕ^* , it can be proved that the secrecy throughput $\mathcal{T}_A(\eta)$ is a concave function of the secrecy rate R_s as done in Theorem 2. Hence, the optimal R_s^* maximizing $\mathcal{T}_A(\eta)$ is given by (12) and a closed-form lower bound on R_s^* can be found in (14). Eventually, problem (26) can be addressed via an alternating optimization (AO) method, which is summarized in Algorithm 1. In addition, at the high η regime, the optimal ϕ^* is independent of R_s , and hence a global optimal pair (ϕ^*, R_s^*) is obtained for maximizing $\mathcal{T}_A(\eta)$.

Fig. 6 illustrates the optimal power allocation ϕ^* and the corresponding maximal secrecy throughput $\mathcal{T}_A(\eta)$ for varying secrecy rate R_s . The maximal $\mathcal{T}_A(\eta)$ is concave on R_s , which guarantees the global optimality of the solution and the convergence of the proposed AO algorithm. The optimal

Algorithm 1 AO Algorithm for Solving Problem (26)

- 1: Initialize $k = 1$, $\phi^{(0)} \in [0, 1]$, $R_s^{(0)} \geq 0$, and assign ϵ a sufficiently small positive value, e.g., $\epsilon = 10^{-10}$;
 - 2: Input $\delta \in [0, 1]$, $N \geq 1$, and $P_b, \Gamma_e, \eta = \|\mathbf{h}_b\|^2 > 0$;
 - 3: Calculate μ from (32) and $\mathcal{T}_A^{(0)}(\eta) = R_s^{(0)} p_s^{(0)}$;
 - 4: **if** $\eta < \mu$ **then**
 - 5: $\mathcal{T}_A^{(k)}(\eta) \leftarrow 0$;
 - 6: **else**
 - 7: Update $\phi^{(k)} \leftarrow \phi^{(k-1)}$, $R_s^{(k)} \leftarrow R_s^{(k-1)}$;
 - 8: Calculate ρ_b^o from (38);
 - 9: **if** $P_b \eta \geq \rho_b^o$ **then**
 - 10: $\phi^{(k)} \leftarrow 1$;
 - 11: **else**
 - 12: Calculate $\phi^{(k)}$ from (37);
 - 13: **end if**
 - 14: Calculate $R_s^{(k)}$ from (12);
 - 15: Update $\mathcal{T}_A^{(k)}(\eta) \leftarrow R_s^{(k)} p_s^{(k)}$;
 - 16: **while** $\left| \frac{\mathcal{T}_A^{(k)}(\eta) - \mathcal{T}_A^{(k-1)}(\eta)}{\mathcal{T}_A^{(k-1)}(\eta)} \right| \geq \epsilon$ **do**
 - 17: Update $k \leftarrow k + 1$;
 - 18: Repeat step 7 to step 15;
 - 19: **end while**
 - 20: **end if**
 - 21: Output $\mathcal{T}_A^{(k)}(\eta)$
-

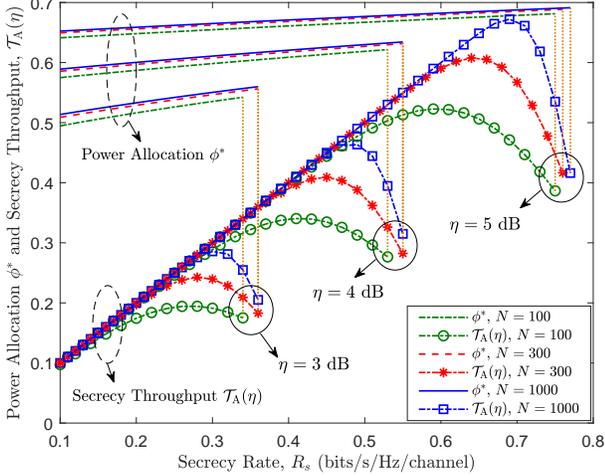


Fig. 6: Optimal ϕ^* and $\mathcal{T}_A(\eta)$ vs. R_s for different N and η , with $M = 4$, $P_b = 0$ dB, $\Gamma_e = 0$ dB, and $\delta = 0.2$.

ϕ^* increases with η and R_s , which verifies Corollary 5. In addition, the curves of ϕ^* and $\mathcal{T}_A(\eta)$ are truncated after R_s exceeds some critical values. This can be explained similarly as that of Fig. 2. It is shown that ϕ^* increases with the blocklength N , although slightly. This is because, increasing N will mildly decrease the information leakage probability \mathcal{O}_e , thus allowing a larger portion of power to be devoted to transmitting the information-bearing signal.

B. Non-Adaptive Optimization Scheme

This subsection examines the secrecy throughput maximization through a non-adaptive design manner for the multi-

antenna transmitter case. The problem can be formulated as

$$\max_{\mu, R_e, R_s, n, \alpha, \phi} \mathcal{T}_N = R_s \bar{p}_s \quad (39a)$$

$$\text{s.t. (6b) - (6d), } 0 \leq \alpha, \phi \leq 1, \quad (39b)$$

where \bar{p}_s is the average successful decoding probability.

The basic idea to solve problem (39) is similar to that of problem (15). Again, the optimal rate redundancy is $R_e^* = \mathcal{O}_e^{-1}(\delta)$ with \mathcal{O}_e given in (4). For the adaptive case, it is known from (27) that $\kappa(\gamma_b(\phi, 1), \alpha) > \kappa(\Phi_e(\phi, 1), \alpha) \Rightarrow \eta > \mu(\phi) = \frac{\rho_e(\phi)}{P_b}$ suffices to guarantee a positive secrecy rate R_s with the threshold $\mu(\phi)$ independent of α , and then $\alpha^* = 1$ is optimal for secrecy throughput maximization. As for the non-adaptive one, supporting a certain secrecy rate R_s requires that $1 + \kappa(\gamma_b(\phi, 1), \alpha) > 2^{R_s} (1 + \kappa(\Phi_e(\phi, 1), \alpha))$ which is further transformed to

$$\eta > \mu(\phi) = \frac{1}{\phi P_b} \frac{1}{\frac{\alpha}{2^{R_s(1+\kappa(\Phi_e(\phi, 1), \alpha))} - 1} + \alpha}. \quad (40)$$

Although $\mu(\phi)$ herein depends on α , it is proved that $\mu(\phi)$ monotonically decreases with α . Hence, $\alpha^* = 1$ is still throughput-optimal for the non-adaptive case. Accordingly, the optimal threshold is $\mu^* = \frac{2^{R_s(1+\phi\rho_e)}}{\phi P_b}$. Similar to the proof of Theorem 3, using the maximal blocklength is optimal for maximizing secrecy throughput, regardless of the distribution of γ_b . Hence, the optimal blocklength is $n^* = N$. Afterwards, the secrecy throughput is calculated from (17):

$$\begin{aligned} \mathcal{T}_N = R_s \bar{p}_s &= R_s \left[1 - \frac{1}{2} \mathcal{F}_{\gamma_b}(\theta_b^2) - \frac{\beta}{\theta_b} \int_{\theta_b^2}^{\tau_b^u} \mathcal{F}_{\gamma_b}(\gamma) d\gamma \right] \\ &\stackrel{(a)}{=} R_s \left[\frac{\bar{\Gamma}(M, \varrho_1)}{2} + \frac{\phi \Gamma_b \beta}{\theta_b} \Delta \Gamma \right], \end{aligned} \quad (41)$$

where (a) holds by invoking the CDF $\mathcal{F}_{\gamma_b}(\gamma)$ of γ_b in (29) and computing the integral, with $\Delta \Gamma \triangleq \sum_{k=0}^{M-1} [\bar{\Gamma}(k+1, \varrho_1) - \bar{\Gamma}(k+1, \varrho_2)]$ and $\bar{\Gamma}(m+1, x) \triangleq \sum_{k=0}^m \frac{x^k e^{-x}}{k!}$ being the regularized upper incomplete gamma function, with $\varrho_1 = \frac{\theta_b^2}{\phi \Gamma_b}$, $\varrho_2 = \frac{\theta_b^2}{\phi \Gamma_b} + \frac{\theta_b}{2\beta \phi \Gamma_b}$, $\theta_b = \sqrt{2^{R_s+R_e^*} - 1}$, and $\beta = \frac{\sqrt{N}}{2\pi}$. Differentiating \mathcal{T}_N w.r.t. ϕ yields

$$\begin{aligned} \frac{d\mathcal{T}_N}{d\phi} &= R_s \left[\frac{\varpi_1 \varrho_1^M e^{-\varrho_1}}{2(M-1)!} + \frac{\phi \Gamma_b \beta \varpi_2 \Delta \Gamma}{\theta_b} + \beta \theta_b \varpi_1 \Gamma(M, \varrho_1) \right. \\ &\quad \left. - \left(\beta \theta_b \varpi_1 + \frac{\varpi_2}{2} \right) \bar{\Gamma}(M, \varrho_2) \right], \end{aligned} \quad (42)$$

where $\varpi_1 = \frac{1}{\phi} - \frac{2^{R_s}}{\theta_b^2} \frac{d\lambda_e}{d\phi}$ and $\varpi_2 = \frac{1}{\phi} - \frac{2^{R_s}}{2\theta_b^2} \frac{d\lambda_e}{d\phi}$ with λ_e given in (34). It is verified that the derivative $\frac{d\mathcal{T}_N}{d\phi}$ is monotonically decreasing with ϕ . In other words, for a fixed R_s , the optimal ϕ^* that maximizes \mathcal{T}_N is unique, which is $\phi^* = 1$ if $\frac{d\mathcal{T}_N}{d\phi}|_{\phi=1} > 0$ or otherwise satisfies $\frac{d\mathcal{T}_N}{d\phi} = 0$. Likewise, it is confirmed that the derivative

$$\begin{aligned} \frac{d\mathcal{T}_N}{dR_s} &= \frac{\bar{\Gamma}(M, \varrho_1)}{2} + \frac{\phi \Gamma_b \beta}{\theta_b} \Delta \Gamma - \frac{\lambda_e R_s 2^{R_s} \ln 2}{\theta_b} \left[\beta \bar{\Gamma}(M, \varrho_1) \right. \\ &\quad \left. + \frac{\phi \Gamma_b \beta \Delta \Gamma}{2\theta_b^2} + \frac{\varrho_1^M e^{-\varrho_1}}{2\theta_b (M-1)!} - \left(\beta + \frac{1}{4\theta_b} \right) \bar{\Gamma}(M, \varrho_2) \right] \end{aligned} \quad (43)$$

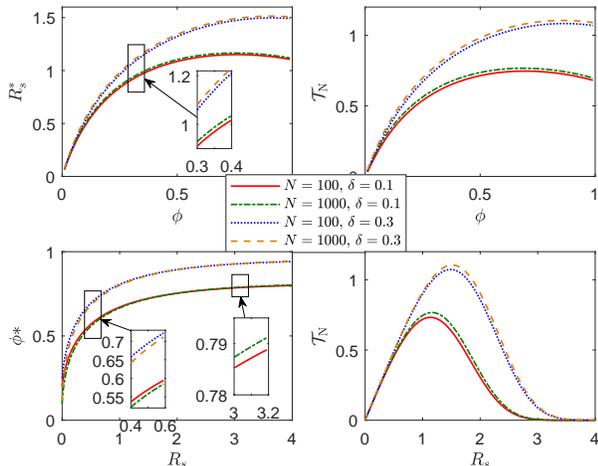


Fig. 7: Above: R_s^* and \mathcal{T}_N vs. ϕ ; Bottom: ϕ^* and \mathcal{T}_N vs. R_s ; for different N and δ , with $M = 4$, $\Gamma_b = 3$ dB and $\Gamma_e = 0$ dB.

is first positive and then negative with increasing R_s , and the unique optimal R_s^* maximizing \mathcal{T}_N can be calculated via a bisection method with the equation $\frac{d\mathcal{T}_N}{dR_s} = 0$.

The monotonicity of \mathcal{T}_N w.r.t. ϕ and R_s is verified in Fig. 7, where, similar to Fig. 6, \mathcal{T}_N is given with the optimal ϕ^* or R_s^* . This implies that the global maximal \mathcal{T}_N is practically achieved even by alternatively solving the optimal ϕ and R_s . As expected, \mathcal{T}_N improves with a larger blocklength N and a looser secrecy constraint (a larger δ). It is found that R_s^* first increases and then might decrease with ϕ , which means that a moderate R_s is desired to balance the decoding and throughput performance. On the other hand, a larger ϕ^* is required to support an increasing R_s . It is also shown that R_s^* for a fixed ϕ increases with N , since a larger N improves the decoding performance which then affords a larger R_s . Nevertheless, ϕ^* decreases with N in the low R_s regime whereas increases with N in the high R_s regime. It can be explained as follows: for a low R_s , the rate redundancy R_e has a great impact on the decoding performance, and hence the AN power should be increased as N increases to better combat the eavesdropper; in contrast, for a large R_s , the decoding correctness is more affected by the main channel quality, which requires a larger signal power to maintain a high decoding probability.

Proposition 1: For the high average channel gain $\Gamma_b \rightarrow \infty$, \mathcal{T}_N in (41) is approximated as

$$\lim_{\Gamma_b \rightarrow \infty} \mathcal{T}_N = R_s \left(1 - \frac{\rho_1^M}{2M!} \right). \quad (44)$$

Proof 14: Please refer to Appendix H.

Proposition 1 shows that for a high average channel gain, the secrecy throughput becomes independent of the blocklength. In consequence, the optimal ϕ^* and R_s^* maximizing \mathcal{T}_N in (44) admit the following closed-form approximations [21], (19),

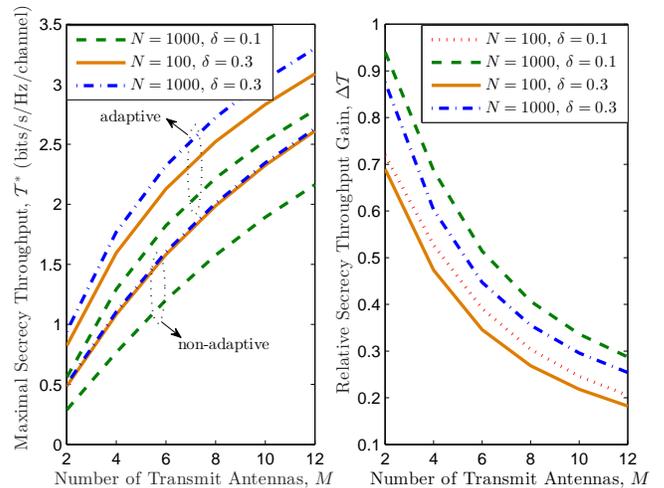


Fig. 8: \mathcal{T}^* and $\Delta\mathcal{T}$ vs. M for different N and δ , with $\Gamma_b = 3$ dB and $\Gamma_e = 0$ dB.

(20)]

$$\lim_{\Gamma_b \rightarrow \infty} \phi^* = \frac{1}{\sqrt{\Lambda} + 1}, \quad (45)$$

$$\lim_{\Gamma_b \rightarrow \infty} R_s^* = \frac{1}{M \ln 2} \left[\mathcal{W}_0 \left(\frac{2 \exp(1) M! \Gamma_b^M}{(\sqrt{\Lambda} + 1)^{2M}} \right) - 1 \right]. \quad (46)$$

Fig. 8 illustrates the influence of the number of transmit antennas M on the maximal secrecy throughput \mathcal{T}^* for both adaptive and non-adaptive schemes and the relative secrecy throughput gain $\Delta\mathcal{T} = \frac{\mathcal{T}_A^* - \mathcal{T}_N^*}{\mathcal{T}_N^*}$. It is not surprising that deploying more transmit antennas can significantly improve the secrecy throughput for both schemes. Similar to the observation in Fig. 4, both \mathcal{T}_A^* and \mathcal{T}_N^* increase with δ and N , but the benefit to \mathcal{T}_N^* brought by a larger N is nearly negligible. The right-hand-side subgraph shows that $\Delta\mathcal{T}$ drops sharply as M increases but grows for a larger N and a smaller δ . This indicates that the superiority of the adaptive scheme over its non-adaptive counterpart is more pronounced for the scenarios requiring a large blocklength, having few transmit antennas, suffering from a stringent secrecy constraint, etc; otherwise, the non-adaptive scheme might be appealing because of the low-complexity off-line design.

C. A Note on Multi-Antenna Eve

This subsection examines the secure transmission in the presence of an Eve with M_e antennas. Assume that Eve employs the minimum mean-squared error (MMSE) receiver, and then the CDF of Eve's SINR under the null-space AN scheme can be given as [23]:

$$\mathcal{F}_{\gamma_e}(x) = 1 - e^{-\frac{x}{\phi P_e}} \sum_{n=1}^{M_e} \frac{A_n(x)}{(n-1)!} \left(\frac{x}{\phi P_e} \right)^{n-1}, \quad (47)$$

where

$$A_n(x) = \begin{cases} 1, & M_e \geq M - 1 + n, \\ \frac{\sum_{m=0}^{M_e-n} \binom{M-1}{m} (\xi x)^m}{(1+\xi x)^{M-1}}, & M_e < M - 1 + n. \end{cases} \quad (48)$$

The information leakage probability \mathcal{O}_e is obtained by substituting (47) into (50), and the secrecy throughput can be optimized similarly as described in the above two subsections.

By ignoring the receiver noise at Eve, i.e., considering Eve's transmit power $P_e \rightarrow \infty$, one can obtain $\mathcal{F}_{\gamma_e}(x) = 1 - A_1(x)$. Furthermore, if Eve has more antennas than Alice, i.e., $M_e \geq M$, one have $A_1(x) = 1$, $\mathcal{F}_{\gamma_e}(x) = 0$, and accordingly $\mathcal{O}_e = 1$. This means, when $P_e \rightarrow \infty$, Eve with enough antennas can completely eliminate all the AN signals with an MMSE receiver such that her SNR will approach infinity. As a consequence, the SOP constraint can no longer be satisfied for any chosen rate redundancy, and no positive secrecy rate can be achieved from the perspective of secrecy outage. In other words, the null-space AN scheme can safeguard secure transmissions well for the finite blocklength regime only when the eavesdropper has fewer antennas than the transmitter, and this conclusion is the same as that for the infinite blocklength case.

V. CONCLUSIONS

This paper investigated the design of secure transmissions in slow fading channels, where secrecy encoding with finite blocklength was employed to confront the eavesdropper. Both adaptive and non-adaptive schemes were devised to maximize the secrecy throughput, providing the optimal threshold of the on-off transmission policy, blocklength, code rates, and power allocation of the AN scheme. Theoretical and numerical results showed that, under the on-off policy, increasing the blocklength can simultaneously enhance the reliability and secrecy, and thus the secrecy throughput is maximized when using the maximal blocklength. In addition, since an overly large secrecy rate will significantly decrease the successful decoding probability thus lowering the secrecy throughput, there exists a critical secrecy rate, but not as large as possible, that can achieve the maximal secrecy throughput.

APPENDIX

A. Proof of Lemma 1

For tractability, a piece-wise linear approximation approach is leveraged to approximate the Q -function given in (2), i.e., $Q\left(\frac{C_i - R_i}{\sqrt{V_i/n}}\right) \approx \Xi(\gamma_i, n, R_i)$ for $i \in \{b, e\}$ [31], [32],³ with

$$\Xi(\gamma_i, n, R_i) = \begin{cases} 0, & \gamma_i > \tau_i^u, \\ \frac{1}{2} - \frac{\beta}{\theta_i}(\gamma_i - \theta_i^2), & \tau_i^l \leq \gamma_i \leq \tau_i^u, \\ 1, & \gamma_i < \tau_i^l, \end{cases} \quad (49)$$

where $\beta \triangleq \frac{\sqrt{n}}{2\pi}$, $\theta_i \triangleq \sqrt{2R_i - 1}$, $\tau_i^u \triangleq \theta_i^2 + \frac{\theta_i}{2\beta}$, and $\tau_i^l \triangleq \theta_i^2 - \frac{\theta_i}{2\beta}$.⁴ With (49), the information leakage probability \mathcal{O}_e defined in (4) is calculated as

$$\mathcal{O}_e = 1 - \mathbb{E}_{\gamma_e} [\Xi(\gamma_e, n, R_e)] = 1 - \frac{\beta}{\theta_e} \int_{\tau_e^l}^{\tau_e^u} \mathcal{F}_{\gamma_e}(\gamma) d\gamma, \quad (50)$$

³ This approximation has been extensively applied to the finite-blocklength scenarios, and its accuracy has been well validated.

⁴ Generally, $\theta_i > \frac{1}{2\beta}$ or $R_i > \log_2(1 + \pi^2/n)$ should be satisfied to ensure a positive τ_i^l .

where $\mathcal{F}_{\gamma_i}(\gamma) = 1 - e^{-\gamma/\Gamma_i^2}$ is the CDF of γ_i for $i \in \{b, e\}$, and the last equality in (50) follows from invoking (49) and using partial integration. Next, treat n as a continuous variable. As R_e^* satisfies $\mathcal{O}_e(R_e^*) = \delta$, the derivative $\frac{dR_e^*}{dn}$ is obtained by using the derivative rule for implicit functions [22] with $\mathcal{O}_e(R_e^*) = \delta$, i.e.:

$$\frac{dR_e^*}{dn} = -\frac{\partial \mathcal{O}_e / \partial n}{\partial \mathcal{O}_e / \partial R_e^*}. \quad (51)$$

First, it can be proved that $\frac{\partial \mathcal{O}_e}{\partial R_e^*} = \frac{\partial \mathcal{O}_e}{\partial \theta_e} \frac{\partial \theta_e}{\partial R_e^*} < 0$ by noting that $\frac{\partial \theta_e}{\partial R_e^*} > 0$ and

$$\begin{aligned} \frac{\partial \mathcal{O}_e}{\partial \theta_e} &= \frac{\beta}{\theta_e^2} \int_{\tau_e^l}^{\tau_e^u} \mathcal{F}_{\gamma_e}(\gamma) d\gamma - \frac{\beta}{\theta_e} \left[\frac{d\tau_e^u}{d\theta_e} \mathcal{F}_{\gamma_e}(\tau_e^u) - \frac{d\tau_e^l}{d\theta_e} \mathcal{F}_{\gamma_e}(\tau_e^l) \right] \\ &\stackrel{(a)}{\leq} \frac{\beta}{\theta_e^2} [\gamma \mathcal{F}_{\gamma_e}(\gamma)] \Big|_{\tau_e^l}^{\tau_e^u} - \frac{4\beta\theta_e + 1}{2\theta_e} \mathcal{F}_{\gamma_e}(\tau_e^u) + \frac{4\beta\theta_e - 1}{2\theta_e} \mathcal{F}_{\gamma_e}(\tau_e^l) \\ &= \beta [\mathcal{F}_{\gamma_e}(\tau_e^l) - \mathcal{F}_{\gamma_e}(\tau_e^u)] < 0, \end{aligned}$$

where (a) follows from the partial integration. The next step is to determine the sign of $\frac{\partial \mathcal{O}_e}{\partial n} = \frac{\partial \mathcal{O}_e}{\partial \beta} \frac{\partial \beta}{\partial n}$. The first and second derivatives of \mathcal{O}_e w.r.t. β are respectively given by

$$\frac{\partial \mathcal{O}_e}{\partial \beta} = \frac{1}{2\beta} [\mathcal{F}_{\gamma_e}(\tau_e^u) + \mathcal{F}_{\gamma_e}(\tau_e^l)] - \frac{1}{\theta_e} \int_{\tau_e^l}^{\tau_e^u} \mathcal{F}_{\gamma_e}(\gamma) d\gamma, \quad (52)$$

$$\frac{\partial^2 \mathcal{O}_e}{\partial \beta^2} = \frac{\theta_e}{4\beta^3} [f_{\gamma_e}(\tau_e^l) - f_{\gamma_e}(\tau_e^u)]. \quad (53)$$

It is easy to see $\frac{\partial^2 \mathcal{O}_e}{\partial \beta^2} > 0$ as $f_{\gamma_e}(\gamma) = \frac{1}{\Gamma_e} e^{-\gamma/\Gamma_e}$ decreases with γ and $\tau_e^u > \tau_e^l$. This indicates that $\frac{\partial \mathcal{O}_e}{\partial \beta}$ increases with β such that $\frac{\partial \mathcal{O}_e}{\partial \beta} < \frac{\partial \mathcal{O}_e}{\partial \beta} \Big|_{\beta \rightarrow \infty} = 0$. Combining $\frac{\partial \mathcal{O}_e}{\partial \beta} < 0$ and $\frac{\partial \beta}{\partial n} > 0$ yields $\frac{\partial \mathcal{O}_e}{\partial n} < 0$. With $\frac{\partial \mathcal{O}_e}{\partial R_e^*} < 0$ and $\frac{\partial \mathcal{O}_e}{\partial n} < 0$ in (51), $\frac{dR_e^*}{dn} < 0$ is obtained, which completes the proof.

B. Proof of Theorem 1

The derivative of p_s w.r.t. n is given by

$$\frac{dp_s}{dn} = \frac{1}{\sqrt{2\pi}} e^{-\frac{n(C_b - R_t)^2}{2V_b}} \left[\frac{C_b - R_t}{2\sqrt{nV_b}} - \sqrt{\frac{n}{V_b}} \frac{dR_e^*}{dn} \right], \quad (54)$$

which follows from the derivative $\frac{dQ(x)}{dx} = \frac{-1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Plugging $\frac{dR_e^*}{dn} < 0$, as shown in Lemma 1, into (54) yields $\frac{dp_s}{dn} > 0$. For a fixed R_s in (10), it is directly concluded that $\frac{dT_A(\eta)}{dn} > 0$, which means that $\mathcal{T}_A(\eta)$ monotonically increases with n . Since n is an integer, $\mathcal{T}_A(\eta)$ is maximized at the maximal integer of n , i.e., $n = N$, which completes the proof.

C. Proof of Theorem 2

From (13), it is easy to prove that $\frac{d^2 \mathcal{T}_A(\eta)}{dR_s^2} < 0$, i.e., $\mathcal{T}_A(\eta)$ is concave on R_s . It is verified that $\frac{d\mathcal{T}_A(\eta)}{dR_s} \Big|_{R_s=0} = 1 - Q\left(\frac{C_b - R_e^*}{\sqrt{V_b/N}}\right) > 0$. As a result, $\mathcal{T}_A(\eta)$ is maximized at the boundary $R_s = C_b - R_e^*$ if $\frac{d\mathcal{T}_A(\eta)}{dR_s} \Big|_{R_s=C_b - R_e^*} = \frac{1}{2} - \frac{C_b - R_e^*}{\sqrt{2\pi V_b/N}} \geq 0 \Rightarrow R_e^* \geq C_b - \sqrt{\frac{\pi V_b}{2N}}$ or otherwise at the unique zero-crossing of $\frac{d\mathcal{T}_A(\eta)}{dR_s}$, i.e., R_s^0 . Next, the condition

$R_e^* \geq C_b - \sqrt{\frac{\pi V_b}{2N}}$ is equivalently transformed to that γ_b does not exceed a critical value γ_b° . Let $\psi(\gamma_b) = C_b - \sqrt{\frac{\pi V_b}{2N}}$. It can be readily confirmed that $\psi(\gamma_b) < 0$, and $\psi(\gamma_b)$ decreases with γ_b if $0 < \gamma_b < \gamma_b^L \triangleq \sqrt{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\pi}{2N}}} - 1$ or otherwise increases with γ_b . This leads to $R_e^* \geq \psi(\gamma_b) \Rightarrow \gamma_b \leq \gamma_b^\circ \triangleq \psi^{-1}(R_e^*)$. An upper bound for γ_b° is further provided by realizing that $\psi(\gamma_b^\circ) = R_e^* > \log_2(1 + \gamma_b^\circ) - \sqrt{\frac{\pi}{2N}} \log_2 e \Rightarrow \gamma_b^\circ < \gamma_b^U \triangleq e^{\sqrt{\frac{\pi}{2N}} + R_e^* \ln 2} - 1$. Then, γ_b° can be quickly calculated using the bisection method with $\psi(\gamma_b) = R_e^*$ in the range (γ_b^L, γ_b^U) . This completes the proof.

D. Proof of Theorem 4

First, display the derivative $\frac{dY(\theta_b)}{d\theta_b}$ in a recursive form $\frac{dY(\theta_b)}{d\theta_b} = \frac{1}{\theta_b} - \left(\frac{1}{\theta_b} + \frac{1}{2\beta\Gamma_b}\right)Y(\theta_b)$ with $Y(\theta_b)$ in (20). Then, the derivative $\frac{dT_N}{d\theta_b}$ is given by $\frac{dT_N}{d\theta_b} = \frac{\theta_b}{1+\theta_b^2} e^{-\frac{\theta_b^2}{\Gamma_b}} G(\theta_b)$, with $G(\theta_b)$ presented in (21). It is easily proved that $G(\theta_b) > 0$ when $\theta_b = \sqrt{2R_e^* - 1}$ and $G(\theta_b) < 0$ as $\theta_b \rightarrow \infty$. The key step of the proof is to argue that $G(\theta_b)$ monotonically decreases with θ_b , which guarantees a unique zero-crossing of $G(\theta_b)$ within $\theta_b \in (\sqrt{2R_e^* - 1}, \infty)$. In other words, T_N initially increases and then decreases with θ_b and reaches the maximum when θ_b arrives at the unique zero-crossing of $G(\theta_b)$. To this end, it is necessary to calculate the derivative $\frac{dG(\theta_b)}{d\theta_b}$:

$$\frac{dG(\theta_b)}{d\theta_b} = \frac{\frac{dY(\theta_b)}{d\theta_b} - 2g(\theta_b)}{\ln 2} - [\log_2(1 + \theta_b^2) - R_e^*] h(\theta_b), \quad (55)$$

where $h(\theta_b) = \left(1 - \frac{1}{\theta_b^2}\right)g(\theta_b) + \frac{1+\theta_b^2}{\theta_b} \frac{dg(\theta_b)}{d\theta_b}$. To proceed, the following lemma is introduced.

Lemma 3: $Y(\theta_b)$ decreases with θ_b and satisfies

$$\frac{2\beta\Gamma_b}{\theta_b + 2\beta\Gamma_b} < Y(\theta_b) < \min\left\{1, \frac{2\beta\Gamma_b}{\theta_b}\right\}. \quad (56)$$

Proof 15: Define $x \triangleq \frac{\theta_b}{2\beta\Gamma_b}$ such that $Y(\theta_b) = \frac{1-e^{-x}}{x}$. The monotonicity of $Y(\theta_b)$ w.r.t. θ_b is due to $\frac{dY(\theta_b)}{d\theta_b} = \frac{(1+x)e^{-x}-1}{x^2} < 0$. The lower bound of $Y(\theta_b)$ is obtained from $\frac{dY(\theta_b)}{d\theta_b} = \frac{1}{\theta_b} - \left(\frac{1}{\theta_b} + \frac{1}{2\beta\Gamma_b}\right)Y(\theta_b) < 0$ and the upper bound follows from $Y(\theta_b) < \frac{1}{x}$ and $1 - e^{-x} < x$.

With the lower bound of $Y(\theta_b)$ given in (56), it can be readily proved that $g(\theta_b) > 0$ such that the term $\frac{dY(\theta_b)}{d\theta_b} - 2g(\theta_b)$ in (55) is negative. Besides, since $h(\theta_b) \geq 0$ directly yields $\frac{dG(\theta_b)}{d\theta_b} < 0$, one only needs to discuss the situation $h(\theta_b) < 0$ and prove that

$$\begin{aligned} \frac{dG(\theta_b)}{d\theta_b} \ln 2 &\leq \frac{dY(\theta_b)}{d\theta_b} - 2g(\theta_b) - h(\theta_b) \ln(1 + \theta_b^2) \\ &\stackrel{(a)}{<} \frac{dY(\theta_b)}{d\theta_b} - 2g(\theta_b) - \theta_b^2 h(\theta_b) \\ &\stackrel{(b)}{\leq} -\frac{1}{\theta_b + 2\beta\Gamma_b} \left(\frac{\theta_b^2}{\Gamma_b} + \frac{\theta_b^4}{\Gamma_b} + 8\beta\theta_b + 8\beta\theta_b^3 - \theta_b^2 \right) \\ &\stackrel{(c)}{<} 0, \end{aligned} \quad (57)$$

where (a) is due to $\ln(1+x) \leq x$, (b) holds by invoking Lemma 3 along with algebraic manipulations, and (c) derives from $8\beta\theta_b + 8\beta\theta_b^3 \geq 16\beta\theta_b^2 > \theta_b^2$ as $\beta = \frac{\sqrt{N}}{2\pi} > \frac{1}{8}$.

E. Proof of Theorem 5

First fix R_s , and it is clear that the term $\frac{-\lambda_b R_s}{\sqrt{\lambda_b^2 - 1}}$ in (28) increases with α as λ_b increases with α . It is also verified that the term $Z(\alpha) \triangleq \frac{\lambda_b(\ln \lambda_b - \ln \lambda_e)}{\sqrt{\lambda_b^2 - 1}}$ in (28) increases with α by computing the derivative of $Z(\alpha)$ w.r.t. α :

$$\begin{aligned} \frac{dZ(\alpha)}{d\alpha} &= \frac{\frac{d\lambda_b}{d\alpha} \left(\lambda_b^2 - 1 - \ln \frac{\lambda_b}{\lambda_e} \right) - \frac{\lambda_b(\lambda_b^2 - 1)}{\lambda_e} \frac{d\lambda_e}{d\alpha}}{(\lambda_b^2 - 1)^{3/2}} \\ &\stackrel{(a)}{=} \frac{\lambda_b(\lambda_b - 1) \left[(\lambda_b - \lambda_e)(\lambda_b + 1) - \ln \frac{\lambda_b}{\lambda_e} \right]}{\alpha(\lambda_b^2 - 1)^{3/2}} \\ &\stackrel{(b)}{\geq} \frac{\lambda_b(\lambda_b - 1)(\lambda_b - \lambda_e) \left(\lambda_b + 1 - \frac{1}{\lambda_e} \right)}{\alpha(\lambda_b^2 - 1)^{3/2}} \stackrel{(c)}{>} 0, \end{aligned} \quad (58)$$

where (a) holds by substituting $\frac{d\lambda_i}{d\alpha} = \frac{\lambda_i(\lambda_i - 1)}{\alpha}$ for $i \in \{b, e\}$, (b) follows from the inequality $\ln \frac{\lambda_b}{\lambda_e} \leq \frac{\lambda_b - \lambda_e}{\lambda_e}$ with $\lambda_b > \lambda_e > 0$, and (c) is due to $\lambda_b > \lambda_e > 1$. Hence, p_s in (28) increases with α as $Q(x)$ decreases with x . This indicates, $\alpha^* = 1$ is optimal for maximizing $\mathcal{T}_A(\eta) = R_s p_s$ for any given R_s and η and is also optimal for maximizing \mathcal{T}_A .

F. Proof of Theorem 7

Let $L(\phi) = L_1 L_2$, where $L_1 = \frac{\lambda_b}{\sqrt{\lambda_b^2 - 1}}$ and $L_2 = \ln \frac{\lambda_b}{\lambda_e} - R_s \ln 2$ such that $\frac{dL_1}{d\phi} = -\rho_b(L_1^2 - 1)^{3/2}$ and $\frac{dL_2}{d\phi} = \frac{\rho_b}{\lambda_b} - \frac{\rho_e + \phi \frac{d\rho_e}{d\phi}}{\lambda_e}$. Rewrite the second derivative as $\frac{d^2 L(\phi)}{d\phi^2} = L_1(L_1^2 - 1)^2 I(\phi)$ with $I(\phi)$ given by (59) at the top of this page, where (a) holds by recalling the definition $L_2 \leq \ln \frac{\lambda_b}{\lambda_e}$ and invoking the result $\frac{d^2 \rho_e}{d\phi^2} > \frac{2}{\rho_e} \left(\frac{d\rho_e}{d\phi} \right)^2 > 0$ from Lemma 2, (b) follows from plugging $L_1 = \frac{\lambda_b}{\sqrt{\lambda_b^2 - 1}}$, using the inequality $\ln \frac{\lambda_b}{\lambda_e} \leq \frac{\lambda_b - \lambda_e}{\lambda_e}$, and omitting the term $(\phi^2 + \frac{2\phi}{\rho_e}) \left(\frac{d\rho_e}{d\phi} \right)^2$, (c) holds by substituting $\frac{dL_2}{d\phi}$ into (b), and (d) is established after some manipulation operations and by discarding the negative term $\frac{2(\lambda_b^2 - 1)}{\lambda_b \lambda_e^2} [\phi \rho_b \lambda_e - \lambda_b (\lambda_b^2 - 1)]$ noting that $\lambda_b = 1 + \phi \rho_b > \lambda_e$. As indicated by (59) that $L(\phi)$ is concave on ϕ , $L(\phi)$ is maximized at $\phi = 1$ if $\frac{dL(\phi)}{d\phi}|_{\phi=1} \geq 0$ or otherwise at the unique zero-crossing of $\frac{dL(\phi)}{d\phi}$. Besides, $\frac{dL(\phi)}{d\phi}|_{\phi=1} \geq 0$ is equivalent to $X(\rho_b) \geq 0$ in (38). Clearly, $A_1 = \frac{(1+\Gamma_e)\rho_e(1)}{1+\rho_e(1)} < 1$ must be ensured to yield a positive $X(\rho_b)$, with which it can be verified that $X(\rho_b)$ increases with ρ_b . As a consequence, $\frac{dL(\phi)}{d\phi}|_{\phi=1} \geq 0$ can be transformed to an explicit form with relation to ρ_b , namely, $\rho_b \geq \rho_b^\circ$. This completes the proof.

G. Proof of Corollary 5

Let $K(\phi)$ denote $\frac{dL(\phi)}{d\phi}$ in (37). It is verified that $\frac{d\phi^*}{d\eta} = -\frac{\partial K(\phi)/\partial \eta}{\partial K(\phi)/\partial \phi}|_{\phi=\phi^*} > 0$ by recalling that $\frac{\partial K(\phi)}{\partial \phi}|_{\phi=\phi^*} < 0$ from

$$\begin{aligned}
I(\phi) &= 3\rho_b^2 L_2 - \frac{2\rho_b \frac{dL_2}{d\phi}}{L_1(L_1^2 - 1)^{1/2}} - \frac{1}{(L_1^2 - 1)^2} \left[\frac{\rho_b^2}{\lambda_b^2} + \frac{2\frac{d\rho_e}{d\phi} + (\phi + \phi^2 \rho_e) \frac{d^2 \rho_e}{d\phi^2} - \rho_e^2 - \phi^2 \left(\frac{d\rho_e}{d\phi}\right)^2}{\lambda_e^2} \right] \\
&\stackrel{(a)}{\leq} 3\rho_b^2 \ln \frac{\lambda_b}{\lambda_e} - \frac{2\rho_b}{L_1 \sqrt{L_1^2 - 1}} \frac{dL_2}{d\phi} - \frac{1}{(L_1^2 - 1)^2} \left[\frac{\rho_b^2}{\lambda_b^2} + \frac{2\frac{d\rho_e}{d\phi} + (\phi^2 + \frac{2\phi}{\rho_e}) \left(\frac{d\rho_e}{d\phi}\right)^2 - \rho_e^2}{\lambda_e^2} \right] \\
&\stackrel{(b)}{\leq} \left[3\rho_b^2 \frac{\lambda_b - \lambda_e}{\lambda_e} - 2\rho_b \frac{\lambda_b^2 - 1}{\lambda_b} \frac{dL_2}{d\phi} - (\lambda_b^2 - 1)^2 \left(\frac{\rho_b^2}{\lambda_b^2} + \frac{2\frac{d\rho_e}{d\phi} - \rho_e^2}{\lambda_e^2} \right) \right] \\
&\stackrel{(c)}{=} \frac{\rho_b^2 [3\lambda_b^2(\lambda_b - \lambda_e) + \lambda_e(1 - \lambda_b^4)]}{\lambda_b^2 \lambda_e} + \frac{2\rho_b(\lambda_b^2 - 1)}{\lambda_b \lambda_e} \left(\rho_e + \phi \frac{d\rho_e}{d\phi} \right) + \frac{(\lambda_b^2 - 1)^2}{\lambda_e^2} \left(\rho_e^2 - 2\frac{d\rho_e}{d\phi} \right) \\
&\stackrel{(d)}{\leq} -\frac{\phi \rho_b^2 (\rho_b - \rho_e)}{\lambda_b^2 \lambda_e^2} [2\phi^4 \rho_b^3 \rho_e + \phi^3 \rho_b^2 (\rho_b + 6\rho_e) + \phi^2 \rho_b (\rho_b + 8\rho_e) + 5\phi \rho_e + 1] < 0, \tag{59}
\end{aligned}$$

Theorem 7 and proving that

REFERENCES

$$\begin{aligned}
\frac{\partial K(\phi)}{\partial \rho_b} \Big|_{\phi=\phi^*} &= \frac{\frac{\lambda_b - \lambda_b + 1}{\lambda_b} - (1 - A_{\phi^*}) + \frac{(2\lambda_b - 1)(\ln \lambda_b - B_{\phi^*})}{\lambda_b + 1}}{\phi^* (\lambda_b^2 - 1)^{5/2} / P_b} \\
&\stackrel{(a)}{=} \frac{\frac{1}{\lambda_b} - \lambda_b + (2\lambda_b^2 - \lambda_b - 1)(1 - A_{\phi^*})}{\phi^* (\lambda_b^2 - 1)^{5/2} / P_b} \\
&\stackrel{(b)}{\geq} \frac{\lambda_b - 1}{\phi^* (\lambda_b^2 - 1)^{5/2} / P_b} > 0, \tag{60}
\end{aligned}$$

where (a) is due to $\frac{\ln \lambda_b - B_{\phi^*}}{\lambda_b + 1} = (1 - A_{\phi^*})\lambda_b - 1$ from $K(\phi^*) = 0$, and (b) is because $(1 - A_{\phi^*})\lambda_b - 1 > 0$. Moreover, $\lim_{\eta \rightarrow \infty} K(\phi^*) = \frac{1 - A_{\phi^*}}{\phi^*}$. Solving $K(\phi^*) = 0$ with $A_{\phi^*} = \frac{\phi^* \Lambda}{(1 - \phi^*)(1 - \phi^* + \phi^* \Lambda)}$ yields $\phi^* = \frac{1}{\sqrt{\Lambda + 1}}$. Similarly, one can prove that $\frac{d\phi^*}{d\Lambda} < 0 \Rightarrow \frac{d\phi^*}{d\delta} > 0$ and $\lim_{\delta \rightarrow 1} \phi^* = 1$.

H. Proof of Proposition 1

Note that $\varrho_1, \varrho_2 \rightarrow 0$ as $\Gamma_b \rightarrow \infty$. Resorting to [21, Eqn. (44)] yields

$$\bar{\Gamma}(M, \varrho_i) = e^{-\varrho_i} \sum_{k=0}^{M-1} \frac{\varrho_i^k}{k!} \approx 1 - \frac{\varrho_i^M}{M!}, \quad i \in \{1, 2\}, \tag{61}$$

and the term $\Delta\Gamma$ in (41) is approximated as

$$\begin{aligned}
\Delta\Gamma &= \sum_{k=0}^{M-1} [\bar{\Gamma}(k+1, \varrho_1) - \bar{\Gamma}(k+1, \varrho_2)] \\
&= \sum_{k=0}^{M-1} \left(e^{-\varrho_1} \sum_{m=0}^k \frac{\varrho_1^m}{m!} - e^{-\varrho_2} \sum_{m=0}^k \frac{\varrho_2^m}{m!} \right) \\
&= \sum_{k=0}^{M-1} \frac{M-k}{k!} (e^{-\varrho_1} \varrho_1^k - e^{-\varrho_2} \varrho_2^k) \\
&= M\Delta\Gamma(M-1, \varrho_1, \varrho_2) - \varrho_1 \bar{\Gamma}(M-1, \varrho_1) + \varrho_2 \bar{\Gamma}(M-1, \varrho_2) \\
&\approx M \left(\frac{\varrho_1^M}{M!} - \frac{\varrho_2^M}{M!} \right) - \left[\varrho_1 - \frac{\varrho_1^M}{(M-1)!} \right] + \left[\varrho_2 - \frac{\varrho_2^M}{(M-1)!} \right] \\
&= \varrho_2 - \varrho_1 = \frac{\theta_b}{2\beta\phi\Gamma_b}. \tag{62}
\end{aligned}$$

Substituting (61) and (62) into (41) completes the proof.

- [1] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.
- [2] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, Aug. 2014.
- [3] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.
- [4] Y. Zou, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.
- [5] H.-M. Wang and T.-X. Zheng, *Physical Layer Security in Random Cellular Networks*. Singapore: Springer, Oct. 2016.
- [6] X. Chen, D. W. K. Ng, W. Gerstacker, and H.-H. Chen, "A survey on multiple-antenna techniques for physical layer security," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1027–1053, 2nd Quart., 2017.
- [7] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and Xiqi Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 679–695, Apr. 2018.
- [8] Y. Deng, L. Wang, S. A. R. Zaidi, J. Yuan, and M. Elkashlan, "Artificial-noise aided secure transmission in large scale spectrum sharing networks," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2116–2129, May 2016.
- [9] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1656–1672, Mar. 2017.
- [10] W. Wang, K. C. Teh, and K. H. Li, "Artificial noise aided physical layer security in multi-antenna small-cell networks," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 6, pp. 1470–1482, Jun. 2017.
- [11] Z. Ding, Z. Zhao, M. Peng, and H. V. Poor, "On the spectral efficiency and security enhancements of NOMA assisted multicast-unicast streaming," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 3151–3163, Jul. 2017.
- [12] S. Yan, N. Yang, I. Land, R. Malaney, J. Yuan, "Three artificial-noise-aided secure transmission schemes in wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3669–3673, Apr. 2018.
- [13] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [14] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.
- [15] V. W. S. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems*. Cambridge, UK: Cambridge University Press, Apr. 2017.
- [16] C. Cao, H. Li, Z. Hu, W. Liu, and X. Zhang, "Physical-layer secrecy performance in finite blocklength case," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec 2015, pp. 1–6.
- [17] W. Yang, R. F. Schaefer, and H. V. Poor, "Finite-blocklength bounds for wiretap channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 3087–3091.

- [18] M. H. Yassaee, M. R. Aref, and A. Gohari, "Non-asymptotic output statistics of random binning and its applications," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 1849–1853.
- [19] V. Y. F. Tan, "Achievable second-order coding rates for the wiretap channel," in *Proc. IEEE Int. Conf. Comm. Syst. (ICCS)*, Singapore, Nov. 2012, pp. 65–69.
- [20] W. Yang, R. F. Schaefer, and H. V. Poor, "Secrecy-reliability tradeoff for semi-deterministic wiretap channels at finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2133–2137.
- [21] 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, Jun. 2013.
- [22] T.-X. Zheng, H.-M. Wang, J. Yuan, D. Towsley, and M. H. Lee, "Multi-antenna transmission with artificial noise against randomly distributed eavesdroppers," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4347–4362, Nov. 2015.
- [23] T.-X. Zheng, H.-M. Wang, Q. Yang, and M. H. Lee, "Safeguarding decentralized wireless networks using full-duplex jamming receivers," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 278–292, Jan. 2017.
- [24] T.-X. Zheng, H.-M. Wang, J. Yuan, Z. Han, and M. H. Lee, "Physical layer security in wireless ad hoc networks under a hybrid full/half-duplex receiver deployment strategy," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3827–3839, Jun. 2017.
- [25] T.-X. Zheng, H.-M. Wang, and J. Yuan, "Secure and energy-efficient transmissions in cache-enabled heterogeneous cellular networks: Performance analysis and optimization," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5554–5567, Nov. 2018.
- [26] J. Farhat, G. Brante, and R. D. Souza, "Secure throughput optimization of selective decode-and-forward with finite blocklength," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Porto, Portugal, Jul. 2018, pp. 1–5.
- [27] W. K. Harrison, J. Almeida, M.R. Bloch, S.W. McLaughlin, and J. Barros, "Coding for secrecy: An overview of error-control coding techniques for physical-layer security," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 41–50, Sep. 2013.
- [28] D. P. Bertsekas and R. G. Gallager, *Data Networks*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice–Hall, 1992.
- [29] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [30] 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, Oct. 2008.
- [31] B. Makki, T. Svensson and M. Zorzi, "Finite block-length analysis of the incremental redundancy HARQ," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 529–532, Oct. 2014.
- [32] B. Makki, T. Svensson, and M. Zorzi, "Wireless energy and information transmission using feedback: Infinite and finite block-length analysis," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5304–5318, Dec. 2016.
- [33] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.
- [34] 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, Jul. 2010.
- [35] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
- [36] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.
- [37] B. Wang, P. Mu, and Z. Li, "Secrecy rate maximization with artificial-noise-aided beamforming for MISO wiretap channels under secrecy outage constraint," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 18–21, Jan. 2015.
- [38] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge, UK: Cambridge University Press, 2010.