Secret Channel Training to Enhance Physical Layer Security With a Full-Duplex Receiver

Shihao Yan, *Member, IEEE*, Xiangyun Zhou, *Senior Member, IEEE*, Nan Yang, *Member, IEEE*, Thushara D. Abhayapala, *Senior Member, IEEE*, and A. Lee Swindlehurst, *Fellow, IEEE*

***Abstract*— This paper proposes a new channel training (CT) scheme for a full-duplex receiver to enhance physical layer security. Equipped with N_B full-duplex antennas, the receiver simultaneously receives the information signal and transmits artificial noise (AN). In order to reduce the non-cancellable self-interference due to the transmitted AN, the receiver has to estimate the self-interference channel prior to the data communication phase. In the proposed CT scheme, the receiver transmits a limited number of pilot symbols that are known only to itself. Such a secret CT scheme prevents an eavesdropper from estimating the jamming channel from the receiver to the eavesdropper, hence effectively degrading the eavesdropping capability. We analytically examine the connection probability (i.e., the probability of the data being successfully decoded by the receiver) of the legitimate channel and the secrecy outage probability due to eavesdropping for the proposed secret CT scheme. Based on our analysis, the optimal power allocation between CT and data/AN transmission at the legitimate transmitter/receiver is determined. Our examination shows that the newly proposed secret CT scheme significantly outperforms the non-secret CT scheme that uses publicly known pilots when the number of antennas at the eavesdropper is larger than one.**

***Index Terms*— Physical layer security, channel training, full duplex, artificial noise, power allocation.**

I. Introduction

As WIRELESS communications become increasingly ubiquitous, a growing amount of research has been devoted to security issues pertaining to wireless data transmission. This is due to the fact that wireless communications are vulnerable to security threats, such as eavesdropping and jamming attacks, due to the open nature of the wireless medium [1], [2]. Against this background, physical layer security is emerging as a promising technique to realize and enhance the secrecy of wireless communications and is also compatible and complementary to traditional cryptographic techniques [3], [4]. For example, the inherent randomness of a wireless channel can be utilized to extract cryptographic keys based on channel reciprocity [5], and physical-layer-based secure communications can be used to distribute keys for the initialization of a wireless network in order to support upper-layer cryptographical techniques.

In the pioneering studies of physical layer security (e.g., [6] and [7]), a wiretap channel was established as the fundamental model to characterize physical layer security, in which an eavesdropper (Eve) attempts to intercept the data transmission between a transmitter (Alice) and a legitimate receiver (Bob). In the context of multiple-input multiple-output (MIMO) wiretap channels, artificial noise (AN)-aided secure transmission is of growing interest due to its robustness and desirable performance (e.g., [8]–[13]). AN was initially proposed to be transmitted in the null space of the main channel (i.e., the channel between Alice and Bob) by Alice to deliberately confuse Eve while avoiding interference to Bob [8]. Then, this AN-aided secure transmission was studied and extended in numerous scenarios. For example, Wang *et al.* [10] proposed an artificial fast fading scheme in order to enhance physical layer security when Eve has more antennas than Alice. The method of transmitting AN by a cooperative jammer (i.e., an external helper) was exploited (e.g., [9]). However, this method suffers from some issues with regard to mobility, synchronization, and trustworthiness [14], [15]. As a result of the full-duplex techniques coming to reality [16]–[18], these issues are being addressed by replacing the external helper by a full-duplex receiver that can simultaneously receive information signals and transmit AN (e.g., [14], [15], and [19]-[22]). Meanwhile, the impact of full-duplex techniques employed by an active eavesdropper on physical layer security was examined within a hierarchical game framework in [23].

One of the key challenges faced in designing practical full-duplex transceivers is self-interference and thus many techniques have been developed in the literature to suppress it [16]–[18]. Among the different types of self-interference cancellation techniques, the channel-aware technique has attracted increasing research interest since it is normally the last line of defense against self-interference

Manuscript received October 18, 2017; revised February 15, 2018 and April 19, 2018; accepted April 25, 2018. Date of publication May 7, 2018; date of current version May 22, 2018. This work was supported by the Australian Research Council's Discovery Projects under Grant DP150103905. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Lifeng Lai. *(Corresponding author: Shihao Yan.)*

S. Yan is with the School of Engineering, Macquarie University, Sydney, NSW 2109, Australia (e-mail: shihao.yan@mq.edu.au).

X. Zhou, N. Yang, and T. D. Abhayapala are with the Research School of Engineering, Australian National University, Canberra, ACT 2601, Australia (e-mail: xiangyun.zhou@anu.edu.au; nan.yang@anu.edu.au; thushara.abhayapala@anu.edu.au).

A. L. Swindlehurst is with the Center for Pervasive Communications and Computing, University of California at Irvine, Irvine, CA 92697 USA (e-mail: swindle@uci.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2018.2834301

in the digital domain [17], [18]. In channel-aware self-interference cancellation, the channel state information (CSI) of the self-interference channel (i.e., the channel between the transmit and receive antennas of Bob) is first estimated and then the self-interference is suppressed by beamforming or subtraction. We note that the self-interference channel in this work refers to the channel from the transmit antenna to the receive antenna or the channel from the transmit RF chain to the receive RF chain when the full-duplex Bob is achieved by connecting two RF chains to a single antenna through a circulator [17]. This is due to the fact that there is leakage in the circulator from the transmit RF chain to the receive RF chain. As such, in this work we do not consider the impact of ADC, DAC, and other hardware impairments on the self-interference cancellation. In the literature, [24] examined a novel secure on-off transmission scheme with AN by considering a practical scenario where the channel training and feedback are limited. It is shown that considering the training and feedback cost there exists an optimal number of transmit antennas that maximizes the net secrecy throughput. In addition, [25] investigated the role of AN in both the channel training and data transmission in physical-layer secret communications, where the AN is used to prevent Eve from estimating the eavesdropper's channel in the training phase and the AN is adopted to mask the transmission of the confidential information in the data transmission phase. However, how to perform the self-interference channel estimation and how to allocate transmit power between the channel training (CT) and data transmission have not been addressed in the context of physical layer security. These questions are of significant importance in the design of practical communication systems to achieve security. This is due to the fact that fewer resources (e.g., transmit power, time slots) are left for data transmission if more resources are allocated to channel estimation (although more accurate channel estimation is achieved). In addition, for channel estimation we have to consider the amount of information about the channels that is leaked to Eve. The secrecy performance of the wiretap channel with a full-duplex receiver is highly related to these questions. The assumption that the CSI of the self-interference channel is perfectly known is widely adopted in the literature and thus the self-interference can be fully cancelled when the full-duplex Bob is equipped with multiple transmit or receive antennas [14]. This assumption cannot be justified in many practical scenarios in which the self-interference channel consists of not only deterministic direct paths but also random reflected paths from nearby scatterers. This partially motivates this work, which, for the first time, examines CT in the wiretap channel with a full-duplex receiver.

In the aforementioned studies where AN is transmitted by a cooperative jammer, the jammer has to know the channel from the jammer to Bob (i.e., jammer-Bob channel) in order to avoid interference to Bob. To this end, public pilots with known transmit power have to be transmitted by the jammer in order to enable the estimation of the jammer-Bob channel at Bob (since they are two separate devices). Meanwhile, Eve can estimate the channel from the jammer to Eve (i.e., jammer-Eve channel) based on the known pilots and transmit power of these pilots. As such, in these studies the jammer-Eve channel is normally assumed to be known by Eve (e.g., [9]). A similar assumption that Eve knows the jamming channel (i.e., the channel from Bob's transmit antennas to Eve) in the wiretap channel with a full-duplex receiver is adopted in the literature (e.g., [14] and [22]). This assumption ignores one property of this wiretap channel, which is that Bob knows exactly the transmitted signals and transmit power. This means that the pilots and transmit power used to estimate the self-interference channel are not required to be public. Ignorance of this property in the literature has meant fact that the benefits of transmitting AN by a full-duplex Bob rather than an external jammer have not been fully exploited. Recently, our conference paper [26] proposed the secret CT design based on this property of the wiretap channel with a full-duplex receiver. However, we would like to clarify that the performance of this proposed design was only examined through simulations in [26], which lead to the fact that the power allocation was also determined based on these simulations. These simulations are of high signal processing cost, which significantly increases the implementation cost of the proposed secret CT design. In order to fully exploit the benefits offered by the secret CT design, in this work we theoretically analyze the performance of this design, which leads to the following specific contributions.

- Following our conference paper [26], in this work we further examine the CT design problem in the context of wiretap channels with a full-duplex receiver. Specifically, we apply the non-secret CT scheme to the wiretap channel with a full-duplex receiver as a benchmark, based on which we further develop a new secret CT scheme by utilizing the fact that a full-duplex Bob knows exactly the signal he transmits.
- In the secret CT scheme, secret pilots are utilized to estimate the self-interference channel with limited symbol periods to prevent Eve from obtaining the CSI of the jamming channel, which is different from the non-secret CT scheme that utilizes publicly known pilots to estimate the self-interference channel. In order to fully explore and analyze the benefits offered by the proposed secret CT scheme, we derive closed-form expressions for its connection probability (CP), which is the probability that Bob successfully decodes Alice's message, and its secrecy outage probability (SOP). Based on the derived CP and SOP, the optimal transmit power allocations between CT and data/AN transmission at Alice and Bob are determined under average power constraints.
- For the sake of comparison, we also examine the secrecy performance of the non-secret CT scheme and the associated power allocation between CT and data/AN transmission in the wiretap channel with a full-duplex receiver. Our examination shows that our proposed secret CT scheme significantly outperforms the non-secret CT scheme when $N_E > 1$, where N_E is the number of antennas at Eve. We also find that the performance advantage of our proposed secret CT scheme increases as N_E increases.

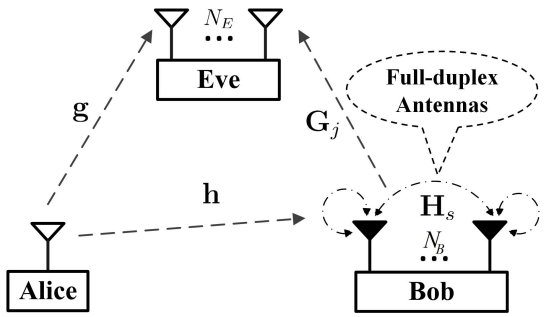


Fig. 1. The wiretap channel with a full-duplex receiver, where Alice is equipped with a single antenna, Bob is equipped with N_B full-duplex antennas, and Eve is equipped with N_E antennas.

The rest of this paper is organized as follows. Section II details the system model, the secret CT scheme, and the non-secret CT scheme. Section III derives the CP and SOP of the secret CT. Section IV presents the optimal transmit power allocations at Alice and Bob in the secret CT scheme. Section V provides numerical results to compare the proposed secret and non-secret CT approaches. Finally, Section VI makes some concluding remarks.

Notation: Scalar variables are denoted by italic symbols. Vectors and matrices are denoted by lower-case and upper-case boldface symbols, respectively. Given a complex number z, $|z|$ denotes the modulus of z. Given a complex vector $\mathbf{x}$, $\|\mathbf{x}\|$ denotes the Euclidean norm and $\mathbf{x}^\dagger$ denotes the conjugate transpose of $\mathbf{x}$. The $L \times L$ identity matrix is referred to as $\mathbf{I}_L$ and $\mathbb{E}[\cdot]$ denotes expectation.

II. System Model

A. Channel Model

The wiretap channel of interest is illustrated in Fig. 1, where Alice is equipped with a single antenna, Bob is equipped with N_B full-duplex antennas, and Eve is equipped with N_E antennas. We assume that Bob operates in full-duplex mode, i.e., all N_B antennas are used for reception and transmission simultaneously. We denote $\mathbf{h} \in \mathbb{C}^{N_B \times 1}$ as the main channel vector, $\mathbf{g} \in \mathbb{C}^{N_E \times 1}$ as the channel vector between Alice and Eve (referred to as the eavesdropper's channel), $\mathbf{G}_j \in \mathbb{C}^{N_E \times N_B}$ as the jamming channel matrix, and $\mathbf{H}_s \in \mathbb{C}^{N_B \times N_B}$ as the self-interference channel matrix. We assume all the wireless channels within our system model are subject to independent quasi-static Rayleigh fading with equal block length. The self-interference channel considered throughout this work is the effective self-interference channel after channel-unaware interference cancellation. Following [16], [18], we know that the deterministic components (e.g., line-of-sight components) in the self-interference channel can be removed through channel-unaware interference cancellation, and thus it is reasonable to assume that the residual self-interference channel after channel-unaware cancellation is subject to independent quasi-static Rayleigh fading. We note that this assumption has been widely adopted in the literature to examine the impact of self-interference in full-duplex systems (e.g., [27]–[29]). We further assume that the entries of $\mathbf{h}$, $\mathbf{g}$, $\mathbf{G}_j$, and $\mathbf{H}_s$ are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with zero-mean. We adopt the assumption that the variance of each entry in $\mathbf{h}$, $\mathbf{g}$, and $\mathbf{G}_j$ is normalized to one, but the variance of each entry in $\mathbf{H}_s$ is σ_s^2. This assumption is to maintain the generality of these channels, since the fading variances (including path loss) of $\mathbf{h}$, $\mathbf{g}$, and $\mathbf{G}_j$ can be effectively absorbed into the noise variance at Bob and the transmit powers of Alice and Bob, while the fading variance of $\mathbf{H}_s$ is quantified by σ_s^2. This is due to the fact that only the statistical distributions of such channels affect the design of the different CT schemes, which will be shown in our following analysis. The assumption that statistical information about $\mathbf{G}_j$ is known can be justified as follows. Considering the independent quasi-static Rayleigh fading, in order to know the distribution of $\mathbf{G}_j$ we only have to know the corresponding large-scale path losses (e.g., the location of Eve). We note that Eve may have been an active transmitter or receiver in previous time slots, where the corresponding large-scale path losses were estimated. Considering the static system settings, we assume the large-scale path losses are fixed or slowly varying, which justifies our assumption that the distribution of $\mathbf{G}_j$ is publicly known. Finally, we note that this assumption is widely adopted in the context of physical layer security (e.g., [30]) and even a stronger assumption that the instantaneous realization of $\mathbf{G}_j$ is known is also widely adopted (e.g., [8], [14], and [31]).

We assume that the total duration of each block consists of T symbol periods, including pilot and data symbols. In the pilot symbol periods, Alice and Bob send pilots to enable the estimation of the main channel and the self-interference channel, respectively. The pilots used by Alice are publicly known. During the data symbol periods, Alice transmits confidential information to Bob while the full-duplex Bob sends AN to aid this secure transmission. We denote Alice's transmit powers for pilots and data by $\mathcal{P}_{Ap}$ and $\mathcal{P}_{Ad}$, respectively. We also denote Bob's transmit powers for pilots and AN by $\mathcal{P}_{Bp}$ and $\mathcal{P}_{Ba}$, respectively. We consider an average power constraint over a fading block [32], in which the total energy for a fading block at a transmitter (i.e., Alice or Bob) is subject to a fixed upper bound. We also consider the passive eavesdropping scenario, in which Alice does not know the CSI of the eavesdropper's channel.

B. Secret Channel Training Scheme

In the considered wiretap channel, the pilots sent by Bob to estimate the self-interference channel can be kept secret from Eve. This is due to the fact that Bob knows exactly what he transmits and thus he does not have to share his pilots with other devices. As such, we next develop a specific CT strategy dedicated to the wiretap channel with a full-duplex receiver, which is named as the secret CT scheme.

In this secret CT method, we first set $T_B = N_B$, where T_B is the number of symbol periods used to estimate the self-interference channel $\mathbf{H}_s$. This assumption is to guarantee a reliable estimate of $\mathbf{H}_s$ at Bob according to the principle of the Linear Minimum Mean Square Error (LMMSE) estimation (i.e., if $T_B < N_B$ Bob cannot achieve a reliable estimate of $\mathbf{H}_s$) [33]. We note that $T_B = N_B$ is also a strict requirement for the secret CT scheme, since when $T_B > N_B$, Eve can

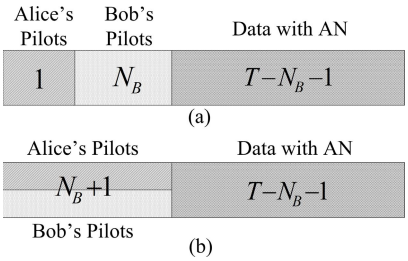


Fig. 2. A transmission block of T symbols with the secret CT scheme and the non-secret CT scheme. (a) Secret channel training scheme. (b) Non-secret channel training scheme.

obtain partial information about the jamming channel $\mathbf{G}_j$ through blind channel estimation [34]. This is due to the fact that once $T_B > N_B$, Eve will have more observations than unknown parameters to estimate and thus she can apply subspace-based channel estimation without knowing the pilots sent by Bob. Setting $T_B = N_B$ guarantees that the estimation problem of $\mathbf{G}_j$ at Eve is ill-posed due to the unknown pilots, and thus Eve cannot achieve any information about $\mathbf{G}_j$ in the secret CT scheme.

To enable Bob to estimate the main channel, Alice transmits its publicly known pilots. We note that Alice and Bob have to transmit pilots during different symbol periods in order to achieve orthogonality between Alice's and Bob's pilots, due to the constraint $T_B = N_B$. In this work, we set the number of symbol periods used to estimate the main channel to be 1 since Alice is equipped with a single antenna. We note that prior studies on optimal training resource allocation have shown that the optimal number of pilot equals the number of transmit antennas (which is N_B for the self-interference channel and 1 for the main channel in this work), under the average power constraint [32]. A transmission block of T symbols with the secret CT scheme is illustrated in Fig. 2 (a).

When Alice transmits the pilot, the corresponding received signal at Bob is given by

$$\mathbf{z}_B = \sqrt{\mathcal{P}_{Ap}}\mathbf{h}s_A + \mathbf{w}_B, \tag{1}$$

where $\mathbf{z}_B \in \mathcal{C}^{N_B\times 1}$, $s_A \in \mathcal{C}^{1\times 1}$ is the pilot transmitted by Alice satisfying $s_A s_A^\dagger = 1$, and $\mathbf{w}_B \in \mathcal{C}^{N_B\times 1}$ is the noise at Bob with i.i.d entries, each of which follows the distribution $\mathcal{CN}(0,\sigma_B^2)$. Considering the LMMSE estimator, based on the known pilot Bob achieves the estimate of $\mathbf{h}$ as [33]

$$\hat{\mathbf{h}} = \frac{\sqrt{\mathcal{P}_{Ap}}}{\mathcal{P}_{Ap}+\sigma_B^2}\mathbf{z}_B s_A^\dagger. \tag{2}$$

Based on the properties of LMMSE estimation [33], the entries of $\hat{\mathbf{h}}$ are i.i.d and each follows the distribution $\mathcal{CN}(0,\sigma_{\hat{h}}^2)$, where

$$\sigma_{\hat{h}}^2 = \frac{\mathcal{P}_{Ap}}{\mathcal{P}_{Ap}+\sigma_B^2}. \tag{3}$$

Again, due to the properties of LMMSE estimation, the estimation error $\tilde{\mathbf{h}} = \mathbf{h} - \hat{\mathbf{h}}$ is independent of $\hat{\mathbf{h}}$ and the entries of $\tilde{\mathbf{h}}$ are i.i.d, each of which follows the distribution $\mathcal{CN}(0,\sigma_{\tilde{h}}^2)$, where

$$\sigma_{\tilde{h}}^2 = \frac{\sigma_B^2}{\mathcal{P}_{Ap}+\sigma_B^2}. \tag{4}$$

Since Alice's pilot is publicly known, Eve can estimate the eavesdropper's channel $\mathbf{g}$ following a similar procedure as detailed above. Likewise, the entries of her estimate on $\mathbf{g}$, denoted by $\hat{\mathbf{g}}$, are i.i.d and each of them follows the distribution $\mathcal{CN}(0,\sigma_{\hat{g}}^2)$, where

$$\sigma_{\hat{g}}^2 = \frac{\mathcal{P}_{Ap}}{\mathcal{P}_{Ap}+\sigma_E^2}, \tag{5}$$

and σ_E^2 is the receiver noise power at Eve. The estimation error $\tilde{\mathbf{g}} = \mathbf{g} - \hat{\mathbf{g}}$ is independent of $\hat{\mathbf{g}}$ and the entries of $\tilde{\mathbf{g}}$ are i.i.d, each of which follows the distribution $\mathcal{CN}(0,\sigma_{\tilde{g}}^2)$, where

$$\sigma_{\tilde{g}}^2 = \frac{\sigma_E^2}{\mathcal{P}_{Ap}+\sigma_E^2}. \tag{6}$$

When Bob transmits pilots over N_B symbol periods with his N_B full-duplex antennas, the signal at his receive antennas is given by

$$\mathbf{Z}_B = \sqrt{\mathcal{P}_{Bp}}\mathbf{H}_s\mathbf{S}_B + \mathbf{W}_B, \tag{7}$$

where $\mathbf{Z}_B \in \mathcal{C}^{N_B\times N_B}$, $\mathbf{S}_B \in \mathcal{C}^{N_B\times N_B}$ are the pilots transmitted by Bob satisfying $\mathbf{S}_B\mathbf{S}_B^\dagger = \mathbf{I}_{N_B}$, and $\mathbf{W}_B \in \mathcal{C}^{N_B\times N_B}$ is the noise at Bob with i.i.d entries, each of which follows the distribution $\mathcal{CN}(0,\sigma_B^2)$. Again, adopting the LMMSE estimator (based on the known $\mathbf{S}_B$ and σ_s^2) Bob obtains the estimate of $\mathbf{H}_s$ as

$$\hat{\mathbf{H}}_s = \frac{\sqrt{\mathcal{P}_{Bp}}\sigma_s^2}{\mathcal{P}_{Bp}\sigma_s^2+\sigma_B^2}\mathbf{Z}_B\mathbf{S}_B^\dagger. \tag{8}$$

Likewise, the estimation error $\tilde{\mathbf{H}}_s = \mathbf{H}_s - \hat{\mathbf{H}}_s$ is independent of $\hat{\mathbf{H}}_s$ and each of its entries follows the distribution $\mathcal{CN}(0,\sigma_{\tilde{H}_s}^2)$, where

$$\sigma_{\tilde{H}_s}^2 = \frac{\sigma_B^2\sigma_s^2}{\mathcal{P}_{Bp}\sigma_s^2+\sigma_B^2}. \tag{9}$$

When Bob transmits the pilots $\mathbf{S}_B$, the received signal matrix at Eve is given by

$$\mathbf{Z}_E = \sqrt{\mathcal{P}_{Bp}}\mathbf{G}_j\mathbf{S}_B + \mathbf{W}_E, \tag{10}$$

where $\mathbf{W}_E \in \mathcal{C}^{N_E\times N_B}$ is the noise at Eve with i.i.d entries, each of which follows the distribution $\mathcal{CN}(0,\sigma_E^2)$. Due to $\mathbf{Z}_E \in \mathcal{C}^{N_E\times N_B}$ and that Eve does not know the pilots $\mathbf{S}_B$, Eve cannot achieve any information on $\mathbf{G}_j$. We note that in order to prevent Eve from obtaining any information on $\mathbf{G}_j$, the communication system should be carefully designed. For example, Eve could learn $\mathbf{G}_j$ based on the control messages (used to establish the communication link, e.g., synchronization) or feedback (e.g., used to feed back the CSI of the main channel to Alice if Alice is equipped with multiple antennas) sent from Bob to Alice. To prevent this, different resources (e.g., frequencies) must be used for control messages or feedback than those used for data communications. In this work, we assume that this independence can be guaranteed and Eve

can only learn statistical information (e.g., the distribution) of $\mathbf{G}_j$ based on pre-existing transmissions between Alice and Bob.

We note that in the context of wireless energy transfer (WET), a channel learning method that requires only one feedback bit for each energy receiver was proposed in [37], which achieves a higher energy transfer efficiency. Relative to this channel learning method, our proposed channel training scheme targets at improving physical layer security in the context of wiretap channels. We do not consider the cost of feedback in the proposed channel training scheme, since in the considered system the transmitter is only equipped with one antenna and multiple antennas are only considered in the full-duplex receiver. Our proposed scheme utilizes one property of the full-duplex receiver to keep the pilot sequences unknown to the eavesdropper (but known at the receiver) in order to prevent the eavesdropper from learning her corresponding channels, which leads to an improved secrecy performance.

C. Data Transmission With AN Following Secret CT

During the data symbol periods, Alice transmits a data stream while Bob transmits AN to confuse Eve. In addition to Eve, the AN also causes interference to Bob through the self-interference channel due to channel estimation errors. In general, Bob has two strategies to suppress such interference based on the estimated self-interference channel $\hat{\mathbf{H}}_s$. First, Bob can subtract the known part of AN based on $\hat{\mathbf{H}}_s$ at his receive antennas since Bob knows the AN he transmits. Second, Bob can transmit AN in the null space of $\hat{\mathbf{H}}_s$, based on the idea that AN that lies in the null space of $\hat{\mathbf{H}}_s$ does not cause any interference to Bob. We note that the second approach requires that the number of Bob's transmit antennas is greater than that of his receive antennas, which is not satisfied in our system model. Therefore, in this work we assume that Bob adopts the first strategy to suppress the interference caused by the AN.

The received signal at Bob in each data symbol period is given by

$$\mathbf{y}_B = \sqrt{\mathcal{P}_{Ad}}\mathbf{h}x + \sqrt{\frac{\mathcal{P}_{Ba}}{N_B}}\mathbf{H}_s\mathbf{n} + \mathbf{v}_B, \tag{11}$$

$$= \sqrt{\mathcal{P}_{Ad}}(\hat{\mathbf{h}} + \tilde{\mathbf{h}})x + \sqrt{\frac{\mathcal{P}_{Ba}}{N_B}}(\hat{\mathbf{H}}_s + \tilde{\mathbf{H}}_s)\mathbf{n} + \mathbf{v}_B, \tag{12}$$

where $x \in \mathcal{C}^{1\times 1}$ denotes the transmitted signal satisfying $\mathbb{E}[|x|^2] = 1$, $\mathbf{n} \in \mathcal{C}^{N_B\times 1}$ is the AN vector, whose entries are i.i.d circularly-symmetric complex normal random variables with zero mean and unit variance, and $\mathbf{v}_B \in \mathcal{C}^{N_B\times 1}$ is the noise vector at Bob with i.i.d entries, each of which follows the distribution $\mathcal{CN}(0, \sigma_B^2)$. Knowing $\hat{\mathbf{H}}_s$ and $\mathbf{n}$, Bob can remove $\hat{\mathbf{H}}_s\mathbf{n}$ from $\mathbf{y}_B$ by subtraction and obtain the effective received signal as

$$\mathbf{y}'_B = \sqrt{\mathcal{P}_{Ad}}\hat{\mathbf{h}}x + \sqrt{\mathcal{P}_{Ad}}\tilde{\mathbf{h}}x + \sqrt{\frac{\mathcal{P}_{Ba}}{N_B}}\tilde{\mathbf{H}}_s\mathbf{n} + \mathbf{v}_B, \tag{13}$$

where $\sqrt{\mathcal{P}_{Ba}/N_B}\tilde{\mathbf{H}}_s\mathbf{n}$ is the residual self-interference. Although Bob knows that his received signal is subject to interference caused by channel estimation errors in $\mathbf{h}$ and $\mathbf{H}_s$, he cannot suppress such interference since he does not know $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{H}}_s$. As such, the optimal combining technique that maximizes the signal-to-interference-plus-noise ratio (SINR) at Bob is maximum ratio combining (MRC) based on $\hat{\mathbf{h}}$. Then, focusing on the CP (similar to the outage probability) and following [38], the instantaneous SINR at Bob for given $\hat{\mathbf{h}}$, $\tilde{\mathbf{h}}$, and $\tilde{\mathbf{H}}_s$ is

$$\gamma_B = \frac{\mu_B\|\hat{\mathbf{h}}\|^2}{\frac{\mu_B|\hat{\mathbf{h}}^\dagger\tilde{\mathbf{h}}|^2}{\|\hat{\mathbf{h}}\|^2} + \frac{\mu_S\|\hat{\mathbf{h}}^\dagger\tilde{\mathbf{H}}_s\|^2}{\|\hat{\mathbf{h}}\|^2} + 1}, \tag{14}$$

where $\mu_B = \mathcal{P}_{Ad}/\sigma_B^2$ and $\mu_S = \mathcal{P}_{Ba}/\sigma_B^2/N_B$.

Likewise, the received signal at Eve in one data symbol period is given by

$$\mathbf{y}_E = \sqrt{\mathcal{P}_{Ad}}\hat{\mathbf{g}}x + \sqrt{\mathcal{P}_{Ad}}\tilde{\mathbf{g}}x + \sqrt{\frac{\mathcal{P}_{Ba}}{N_B}}\mathbf{G}_j\mathbf{n} + \mathbf{v}_E, \tag{15}$$

where $\mathbf{v}_E \in \mathcal{C}^{N_E\times 1}$ is the noise vector at Eve with i.i.d entries, each of which follows the distribution $\mathcal{CN}(0, \sigma_E^2)$. We note that, due to the known structure of the pilots $\mathbf{S}_B$, Eve may still obtain some information on the jamming channel matrix $\mathbf{G}_j$, even though she does not know either $\mathbf{S}_B$ or P_{Bp}. For example, she can approximate $\mathbf{G}_j\mathbf{G}_j^\dagger$ by $\mathbf{Z}_E\mathbf{Z}_E^\dagger$. Due to the uncertainty in the information on $\mathbf{G}_j$ leaked to Eve, it is very difficult to determine the optimal combining strategy at Eve, and the optimal strategy will lead to mathematically intractable analysis. As confirmed numerically, the MRC strategy can outperform the MMSE combining strategy with $\mathbf{G}_j\mathbf{G}_j^\dagger \approx \mathbf{Z}_E\mathbf{Z}_E^\dagger$ in terms of achieving higher SINRs at Eve. As such, in this work we assume that Eve adopts MRC based on $\hat{\mathbf{g}}$ to combine the received signals. Then, following (15) the SINR at Eve for the secret CT scheme is given by

$$\gamma_E = \frac{\mu_E\|\hat{\mathbf{g}}\|^2}{\frac{\mu_E|\hat{\mathbf{g}}^\dagger\tilde{\mathbf{g}}|^2}{\|\hat{\mathbf{g}}\|^2} + \frac{\mu_J\|\hat{\mathbf{g}}^\dagger\mathbf{G}_j\|^2}{\|\hat{\mathbf{g}}\|^2} + 1}, \tag{16}$$

where $\mu_E = \mathcal{P}_{Ad}/\sigma_E^2$ and $\mu_J = \mathcal{P}_{Ba}/\sigma_E^2/N_B$.

D. Traditional Channel Training Scheme as a Benchmark

In order to better illustrate the benefits of the secret CT scheme, we now consider the traditional CT scheme as a benchmark. Unlike the secret CT scheme, in the traditional CT scheme the pilot transmitted by Bob (i.e., $\mathbf{S}_B$) is publicly known, which can be jointly designed with the pilot transmitted by Alice (i.e., s_A). As such, in the traditional CT scheme we do not need the constraint $T_B = N_B$ because Bob's pilots are known by Eve anyway. Hence, Alice and Bob can simultaneously transmit pilots over $1 + N_B$ symbol periods while still ensuring the orthogonality of their pilots. This setting also guarantees a fair comparison between the secret CT scheme and the traditional CT scheme, since the total number of symbol periods allocated to CT is $1 + N_B$ in both schemes. A transmission block of T symbols with the traditional CT scheme is illustrated in Fig. 2 (b). Therefore, $\sigma_{\hat{h}}^2$, $\sigma_{\tilde{h}}^2$, and $\sigma_{\tilde{H}_s}^2$ (which are given by (3), (4), (5), (6), and (9),

respectively, in the secret CT scheme) for the traditional CT scheme are changed to

$$\sigma_{\hat{h}}^2 = \frac{\mathcal{P}_{Ap}(1+N_B)}{\mathcal{P}_{Ap}(1+N_B)+\sigma_B^2}, \tag{17}$$

$$\sigma_{\tilde{h}}^2 = \frac{\sigma_B^2}{\mathcal{P}_{Ap}(1+N_B)+\sigma_B^2}, \tag{18}$$

$$\sigma_{\hat{g}}^2 = \frac{\mathcal{P}_{Ap}(1+N_B)}{\mathcal{P}_{Ap}(1+N_B)+\sigma_E^2}, \tag{19}$$

$$\sigma_{\tilde{g}}^2 = \frac{\sigma_E^2}{\mathcal{P}_{Ap}(1+N_B)+\sigma_E^2}, \tag{20}$$

$$\sigma_{\tilde{H}_s}^2 = \frac{\sigma_B^2\sigma_s^2}{\mathcal{P}_{Bp}\sigma_s^2(1+N_B)/N_B+\sigma_B^2}. \tag{21}$$

In the traditional CT scheme, the pilot $\mathbf{S}_B$ is public and thus Eve can estimate the jamming channel $\mathbf{G}_j$. The estimation error of $\mathbf{G}_j$ can be obtained in a similar format as given in (21). Since Eve knows that her received signal is subject to the interference caused by the channel estimation error and AN, the optimal combining technique that maximizes the SINR at Eve is MMSE based on $\hat{\mathbf{g}}$ and her estimate of $\mathbf{G}_j$ (denoted by $\hat{\mathbf{G}}_j$). Here, we note that MMSE outperforms MRC in terms of achieving a higher SINR at Eve in the traditional CT scheme. This is due to the fact that MMSE utilizes Eve's knowledge of $\hat{\mathbf{G}}_j$ (in addition to $\hat{\mathbf{g}}$) to suppress the interference in order to maximize the SINR, while MRC ignores the available $\hat{\mathbf{G}}_j$. Following (15) and applying the MMSE combiner, for the traditional CT scheme the instantaneous SINR at Eve is

$$\gamma_E = \frac{\mathcal{P}_{Ad}\mathbf{w}\hat{\mathbf{g}}\hat{\mathbf{g}}^\dagger\mathbf{w}^\dagger}{\mathbf{w}\left(\frac{\mathcal{P}_{Ba}}{N_B}\hat{\mathbf{G}}_j\hat{\mathbf{G}}_j^\dagger+\frac{\mathcal{P}_{Ba}}{N_B}\tilde{\mathbf{G}}_j\tilde{\mathbf{G}}_j^\dagger+\mathcal{P}_{Ad}\tilde{\mathbf{g}}\tilde{\mathbf{g}}^\dagger+\sigma_E^2\mathbf{I}_{N_E}\right)\mathbf{w}^\dagger}, \tag{22}$$

where

$$\mathbf{w} = \hat{\mathbf{g}}^\dagger\left(\hat{\mathbf{G}}_j\hat{\mathbf{G}}_j^\dagger+\sigma_E^2\mathbf{I}_{N_E}\right)^{-1}. \tag{23}$$

III. Secrecy Performance Analysis of the Secret Channel Training Scheme

During the data symbol periods, Alice adopts a fixed-rate wiretap code that can be described by two rate parameters, namely, the codeword rate R_B and the redundancy rate R_E, which are predetermined and fixed [40]–[42]. The actual information rate is given by $R_s = R_B - R_E$. For such a coding scheme, Bob cannot reliably decode the transmitted information when the capacity of the main channel (i.e., C_B) is less than R_B, while perfect secrecy against Eve fails when the capacity of the eavesdropper's channel (i.e., C_E) is larger than R_E [40]–[42]. We refer to the probability of achieving reliable decoding as CP and refer to the probability of failing to achieve perfect secrecy as SOP. We note that one conventional secrecy outage probability in the literature is defined as the probability of the secrecy capacity C_s being less than the secrecy rate R_s, where $C_s = C_B - C_E$. In this work, the reliability cannot be guaranteed due to the fact that the main channel suffers from estimation errors, i.e., C_B is not available. As such, if we adopt the conventional secrecy outage probability in this work, it will be a combination of the CP and SOP and cannot separate the CP and SOP. Therefore, in this work we adopt R_B and R_E as the two interesting code parameters to derive the CP and SOP, respectively. Although the ergodic secrecy capacity can be formulated for a passive eavesdropping scenario where Alice and Bob do not know the CSI of the eavesdropper"s channel, in this work we did not adopt it as a performance metric. First, in this work we assume that all wireless channels are subject to independent quasi-static Rayleigh fading, for which outage probability is more appropriate than ergodic channel capacity, regardless of whether secrecy or non-secrecy communication is considered [43]. In addition, there exists a certain SOP associated with the ergodic secrecy capacity. Thus, the ergodic secrecy capacity cannot fully capture the secrecy performance of the system. We also note that the secrecy rate R_s is different from the secrecy capacity examined in [44] from the perspective of information theory. In [44], the secrecy capacity is achieved by encryption over the channel. Without encryption, the secrecy capacity is achieved only when the instantaneous CSI of both the main channel and the eavesdropper's channel is available at Alice. Due to channel estimation error on the main channel and the passive eavesdropping scenario for the eavesdropper's channel, the secrecy transmission rate suffers from connection outages and secrecy outages. As such, in this work we focus on analyzing the CP and SOP.

A. Connection Probability

The CP, which is the probability that Bob can decode the message for a given R_B with a negligible decoding error probability, is given by [38] and [45]

$$P_c = \Pr(\log_2(1+\gamma_B) \geq R_B). \tag{24}$$

We derive the CP of the secret CT scheme in the following theorem.

Theorem 1: The CP of the secret CT scheme is derived as

$$\begin{aligned} P_c = {} & \frac{c^{N_B}e^{-\frac{c}{\mu_h}}}{\Gamma(N_B+1)\Gamma(N_B)(c\mu_H+\mu_h)^{N_B}} \\ & \times\sum_{i=0}^{N_B-1}\frac{\binom{N_B-1}{i}\Gamma(N_B+i+1)}{(c\mu_H+\mu_h)^{i+1}(\mu_H\mu_h)^{-i-1}} \\ & \times\left[{}_2F_1\left(1,N_B+i+1;N_B+1;\frac{\mu_h}{c\mu_H+\mu_h}\right)\right. \\ & \left.-{}_2F_1\left(1,N_B+i+1;N_B+1;\frac{\mu_h(\mu_t-\mu_H)}{\mu_t(c\mu_H+\mu_h)}\right)\right], \end{aligned} \tag{25}$$

where $c = 2^{R_B}-1$, $\mu_h = \mu_B\sigma_{\hat{h}}^2$, $\mu_H = \mu_S\sigma_{\tilde{H}}^2$, $\mu_t = \mu_B\sigma_{\tilde{h}}^2$, $\Gamma(n)$ is the Gamma function given by $\Gamma(n) = (n-1)!$ for positive integer n, and ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ denotes the Gauss hypergeometric function [46, eq. (9.100)].

Proof: The proof is provided in Appendix. ∎

We would like to clarify that the derived CP given in (25) is indeed a function of all power parameters, i.e., $\mathcal{P}_{Ap}$, $\mathcal{P}_{Ad}$, $\mathcal{P}_{Bp}$, and $\mathcal{P}_{Ba}$. This is due to the fact that following (3) and (4) both μ_h and μ_t are functions of $\mathcal{P}_{Ap}$ and $\mathcal{P}_{Ad}$ while μ_H is a function of $\mathcal{P}_{Bp}$ and $\mathcal{P}_{Ba}$ as per (9).

B. Secrecy Outage Probability

The SOP, which is the probability that the capacity of the eavesdropper's channel is no less than R_E, is given by [45]

$$P_{so} = \Pr\left(\log_2(1+\gamma_E) \geq R_E\right). \tag{26}$$

We derive the SOP of the secret CT scheme for positive σ_E^2 in the following theorem.

Theorem 2: The SOP of the secret CT scheme for positive σ_E^2 is derived as

$$\begin{aligned} P_{so} = & \frac{d^{N_E} e^{-\frac{d}{\mu_g}} \mu_g^{N_B-N_E}}{\Gamma(N_B+1)\Gamma(N_E)(d\mu_J+\mu_g)^{N_B}} \\ & \times \sum_{i=0}^{N_E-1} \frac{\binom{N_E-1}{i}\Gamma(N_B+i+1)}{(d\mu_J+\mu_g)^{i+1}(\mu_J\mu_g)^{-i-1}} \\ & \times \left[{}_2F_1\left(1, N_B+i+1; N_B+1; \frac{\mu_g}{d\mu_J+\mu_g}\right)\right. \\ & \left. - {}_2F_1\left(1, N_B+i+1; N_B+1; \frac{\mu_g(\mu_d-\mu_J)}{\mu_d(d\mu_J+\mu_g)}\right)\right], \end{aligned} \tag{27}$$

where $d = 2^{R_E}-1$, $\mu_g = \mu_E\sigma_{\hat{g}}^2$, and $\mu_d = \mu_E\sigma_{\tilde{g}}^2$.

Proof: Comparing (16) with (14), we find that γ_E follows the same type of distribution as γ_B. As such, the proof is similar to the proof of Theorem 1 and thus omitted here. ■

IV. Optimal Transmit Power Allocation in the Secret Channel Training Scheme

We have seen in the previous section that the power values are key design parameters affecting the connection and secrecy performances. In this section, we determine the optimal transmit power allocation between CT and data/AN transmission at Alice and Bob in the secret CT scheme under average power constraints. We note that in the considered passive eavesdropping scenario, the instantaneous SINR at Eve is not available and thus the power allocation does not depend on the instantaneous SINR at Eve. In particular, we cannot achieve the optimal power allocation that maximizes the secrecy capacity of the considered system.

A. Objective Function

As mentioned in Section III, data transmission in the considered wiretap channel may incur connection and secrecy outages. Considering block fading channels, we adopt the effective throughput subject to a given secrecy constraint as our key performance metric, which is given by [42]

$$\begin{aligned} &\eta = \frac{T-N_B-1}{T}(R_B-R_E)P_c, \\ &\text{s.t. } P_{so} \leq \epsilon, \end{aligned} \tag{28}$$

where ϵ is the maximum allowable SOP (i.e., the predetermined secrecy requirement of the system) and $R_B > R_E$ in order to ensure a positive information rate. In this work, we consider fixed-rate transmission, in which R_B and R_E are *a priori* fixed. We note that T is fixed in our system model, and thus the maximization of η subject to $P_{so} \leq \epsilon$ given in (28) is equivalent to the following optimization

$$\max P_c \quad \text{s.t. } P_{so} \leq \epsilon. \tag{29}$$

We note that we cannot guarantee perfect secrecy (i.e., the SOP cannot be zero) in (28) due to the fact that Alice does not have perfect CSI of the eavesdropper's channel due to the passive eavesdropping scenario and/or channel estimation errors. We also note that in this work we cannot consider the case of $N_B \to \infty$ for a given finite T. This is due to the fact that the number of time slots used to estimate the self-interference channel is at least the same as N_B in order to achieve a reliable channel estimate. In order to achieve a positive effective secrecy throughput, we have $N_B < T-1$ as per (28).

B. Power Allocation Under Average Power Constraints

In this work, we consider an average power constraint at both Alice and Bob. Following (29) the power allocation optimization for the secret CT scheme can be presented as

$$\max_{\mathcal{P}_{Ap},\mathcal{P}_{Ad},\mathcal{P}_{Bp},\mathcal{P}_{Ba}} P_c, \tag{30}$$

$$\text{s.t. } P_{so} \leq \epsilon, \tag{31}$$

$$\mathcal{P}_{Ap} + \mathcal{P}_{Ad}(T-N_B-1) \leq \mathcal{E}_A, \tag{32}$$

$$\mathcal{P}_{Bp}N_B + \mathcal{P}_{Ba}(T-N_B-1) \leq \mathcal{E}_B, \tag{33}$$

where $\mathcal{E}_A$ and $\mathcal{E}_B$ are respectively the total powers available at Alice and Bob for each block of T symbol periods; hence, the average power constraints per symbol for Alice and Bob are $\mathcal{E}_A/T$ and $\mathcal{E}_B/T$. We next detail how to determine the solution to (30) (i.e., the optimal values of $\mathcal{P}_{Ad}$, $\mathcal{P}_{Ap}$, $\mathcal{P}_{Ba}$, and $\mathcal{P}_{Bp}$).

Theorem 3: The optimal value of $\mathcal{P}_{Ad}$ that maximizes P_c subject to the constraints given in (31), (32), and (33) can be obtained through

$$\mathcal{P}_{Ad}^* = \underset{0<\mathcal{P}_{Ad}<\mathcal{P}_{Ad}^m}{\operatorname{argmax}} P_c(\mathcal{P}_{Ap}^\dagger, \mathcal{P}_{Ad}, \mathcal{P}_{Bp}^\dagger, \mathcal{P}_{Ba}^\dagger), \tag{34}$$

where

$$\mathcal{P}_{Ad}^m = \frac{\mathcal{E}_A}{T-N_B-1}, \tag{35}$$

$$\mathcal{P}_{Ap}^\dagger = \mathcal{E}_A - \mathcal{P}_{Ad}(T-N_B-1), \tag{36}$$

$$\mathcal{P}_{Bp}^\dagger = \frac{\mathcal{E}_B - \mathcal{P}_{Ba}^\dagger(T-N_B-1)}{N_B}, \tag{37}$$

and $\mathcal{P}_{Ba}^\dagger$ as a function of $\mathcal{P}_{Ad}$ can be obtained by solving the following equation

$$P_{so} = \epsilon. \tag{38}$$

Proof: We first note that P_{so} and P_c are both monotonically decreasing functions of $\mathcal{P}_{Ba}$. As such, $P_{so} = \epsilon$ is always achieved in order to maximize P_c subject to the secrecy constraint (31). Otherwise (i.e., if $P_{so} < \epsilon$), we can decrease $\mathcal{P}_{Ba}$ to increase P_c. Following (27), we note that P_{so} only depends on $\mathcal{P}_{Ba}$ and $\mathcal{P}_{Ad}$. As such, we can obtain $\mathcal{P}_{Ba}^\dagger$ as a function of $\mathcal{P}_{Ad}$ by solving (38). We also note that P_{so} is not a function of $\mathcal{P}_{Ap}$ or $\mathcal{P}_{Bp}$, while P_c monotonically increases as $\mathcal{P}_{Ap}$ or $\mathcal{P}_{Bp}$ increases. Then, we can conclude that the equality in both (32) and (33) is always guaranteed, which leads to (36) and (37), respectively. Finally, (35) is achieved due to $\mathcal{P}_{Ap} > 0$. This completes the proof of Theorem 3. ■

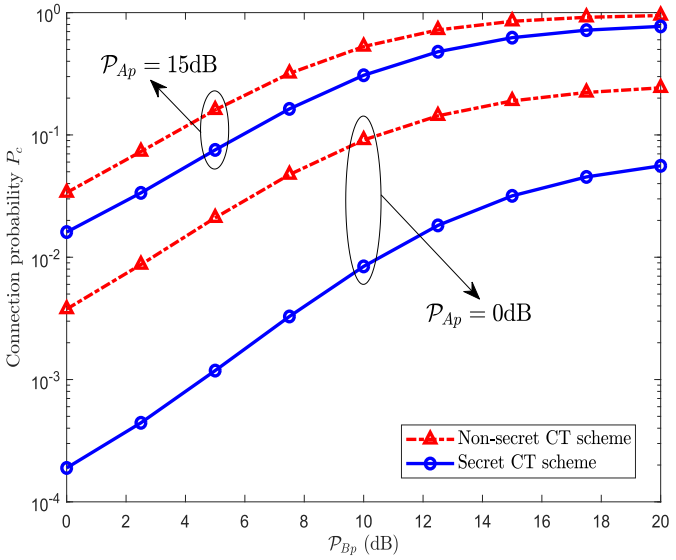


Fig. 3. Connection probability P_c versus $\mathcal{P}_{Bp}$ for different values of $\mathcal{P}_{Ap}$ where $\mathcal{P}_{Ad} = \mathcal{P}_{Ba} = 20$dB, $R_B = 5$, $N_B = 3$, $\sigma_s^2 = 1$, and $\sigma_B^2 = 1$.

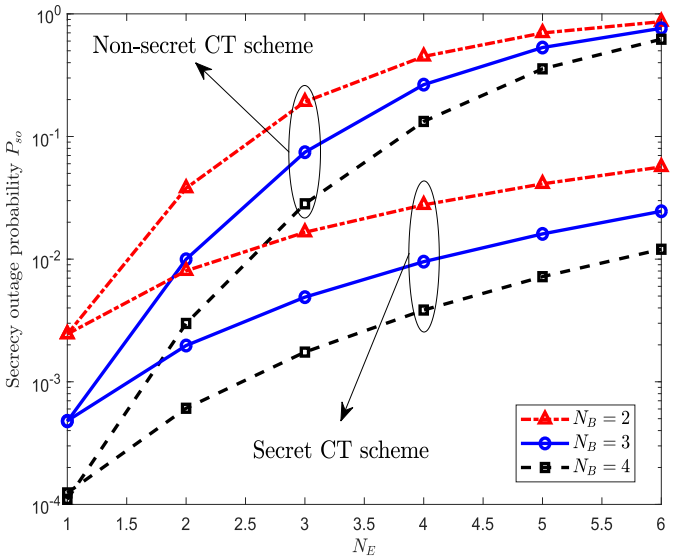


Fig. 4. Secrecy outage probability P_{so} versus N_E for different values of N_B, where $\mathcal{P}_{Ad} = 10$dB, $\mathcal{P}_{Ba} = 20$dB, $\mathcal{P}_{Ap} = 10$dB, $\mathcal{P}_{Bp} = 20$dB, $\sigma_E^2 = 1$, and $R_E = 2$.

By substituting $\mathcal{P}_{Ad}^*$, $\mathcal{P}_{Ap}^*$, $\mathcal{P}_{Ba}^*$, and $\mathcal{P}_{Bp}^*$ into (25), we can obtain the maximum CP of the secret CT scheme. We note that the optimization problem given in (34) can be solved by a one-dimensional numerical search. Specifically, we can adopt a grid search on $\mathcal{P}_{Ad}$ to solve this optimization problem, since its value must lie in the interval $(0, \mathcal{P}_{Ad}^m)$ as given in Theorem 3. We can pick a value of $\mathcal{P}_{Ad}$ within $(0, \mathcal{P}_{Ad}^m)$, then determine the values of $\mathcal{P}_{Ap}^\dagger$, $\mathcal{P}_{Ba}^\dagger$, and $\mathcal{P}_{Bp}^\dagger$ as per Theorem 3, and finally achieve a value of P_c following Theorem 1.

V. Numerical Results

In this section, we present numerical results to examine the secrecy performance of the proposed secret CT scheme with the non-secret CT scheme as the benchmark. We note that it is the ratio between the transmit power (i.e.,$\mathcal{P}_{Ap}$, $\mathcal{P}_{Ad}$, $\mathcal{P}_{Bp}$, and $\mathcal{P}_{Ba}$) and the receiver noise power (i.e., σ_B^2) that affects the numerical results. We also note that the unit of the wiretap code rates (i.e., R_B and R_E) in this work is the bit, which is omitted as well.

In Fig. 3 we plot P_c with different choices of pilot transmit power values. In this figure, we first observe that P_c increases as $\mathcal{P}_{Ap}$ or $\mathcal{P}_{Bp}$ increases. This is due to the fact that as $\mathcal{P}_{Ap}$ and $\mathcal{P}_{Bp}$ increase, the channel estimation errors for the main self-interference channels decrease, respectively. Furthermore, we observe that the CP of the non-secret CT scheme is higher than that of the secret CT scheme under the specific non-optimized and unconstrained settings. This can be explained by the fact that in the non-secret CT scheme both Alice and Bob transmit pilots in $N_B + 1$ symbol periods while in the secret CT scheme Alice and Bob transmit pilots in 1 and N_B symbol periods, respectively, and thus for fixed $\mathcal{P}_{Ap}$ and $\mathcal{P}_{Bp}$ more power is utilized in the non-secret CT scheme. As we will show in Fig. 5, this observation does not hold when the power allocation is optimized under the average power constraints as well as the secrecy outage constraint.

In Fig. 4 we plot P_{so} versus the number of antennas at Eve for different numbers of antennas at Bob. As expected, in this figure we first observe that when $N_E = 1$ the SOP is the same for the secret CT scheme and the non-secret CT scheme, due to the fact that knowing the CSI of the jamming channel in the non-secret CT scheme does not help Eve for $N_E = 1$. For $N_E > 1$, we observe that P_{so} for the secret CT scheme is significantly lower than that for the non-secret CT scheme and the gap increases as N_E increases. This can be explained by the fact that in the secret CT scheme Eve does not know the jamming channel while she does in the non-secret CT scheme and the CSI of the jamming channel offers more information to Eve as N_E increases. Finally, the figure confirms that P_{so} decreases as N_B increases.

In Fig. 5 we plot the maximum CPs of the secret and non-secret CT schemes versus the secrecy constraint indicator ϵ for different values of N_E. In this figure and the following figures, we set a small value -20dB for σ_E^2. In this figure, we first observe that as ϵ increases the maximum CP increases, which demonstrates the tradeoff between the effective throughput and the secrecy constraint. For example, by comparing the values of the maximum CP for $\epsilon = 0.01$ and $\epsilon = 0.15$ we can see a high cost for reducing the maximum CP to achieve secrecy. We also observe that the proposed secret CT scheme significantly outperforms the non-secret CT scheme in terms of achieving a much higher maximum CP. This is due to the fact that the secret CT scheme prevents Eve from obtaining the CSI of the jamming channel. This observation also demonstrates the benefits of transmitting AN by a full-duplex Bob relative to transmitting AN by an external helper. Finally, we observe that the performance gap between these two schemes increases as N_E increases.

Under the same settings as in Fig. 5, we plot Alice's optimal transmit power for data (i.e., $\mathcal{P}_{Ad}^*$) and Bob's optimal transmit power for AN (i.e., $\mathcal{P}_{Ba}^*$) versus ϵ in Fig. 6 and Fig. 7,

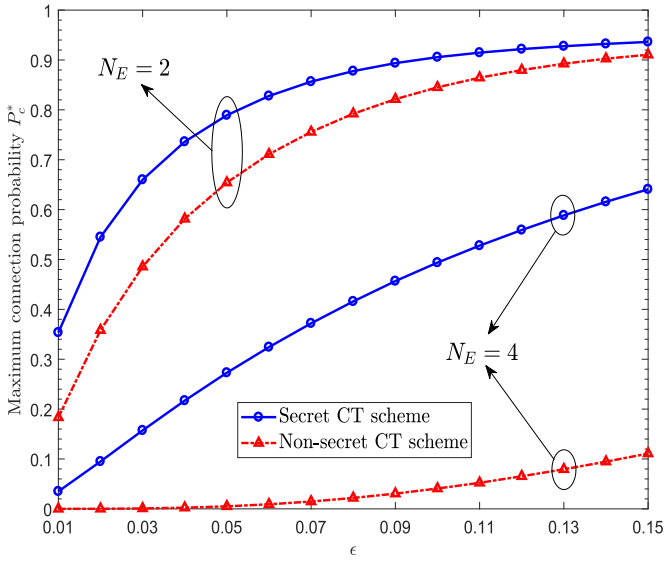


Fig. 5. The maximum connection probability of the secret and non-secret CT schemes versus the secrecy constraint ϵ for different values of N_E, where $N_B = 8$, $R_B = 5$, $R_E = 3$, $T = 300$, $\mathcal{E}_A/T = \mathcal{E}_B/T = 10$dB, $\sigma_s^2 = 1$, and $\sigma_B^2 = 0$dB.

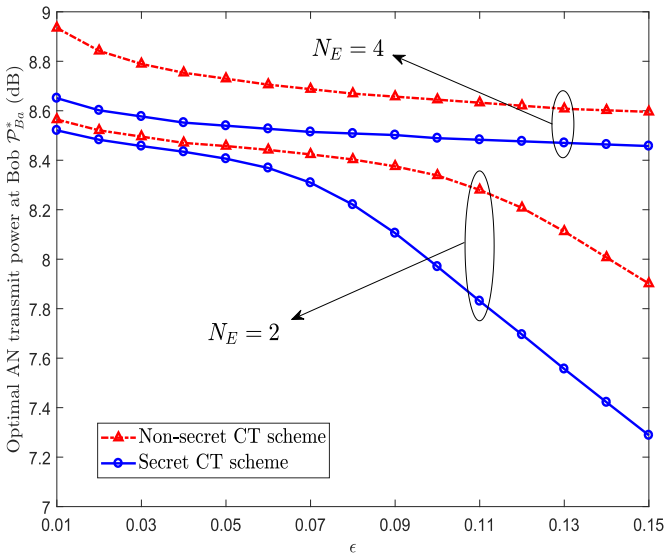


Fig. 7. Bob's optimal AN transmit power $\mathcal{P}_{Ba}^*$ versus the secrecy constraint ϵ for different values of N_E, where $N_B = 8$, $R_B = 5$, $R_E = 3$, $T = 300$, $\sigma_s^2 = 1$, $\mathcal{E}_A/T = \mathcal{E}_B/T = 10$dB, and $\sigma_B^2 = 0$dB.

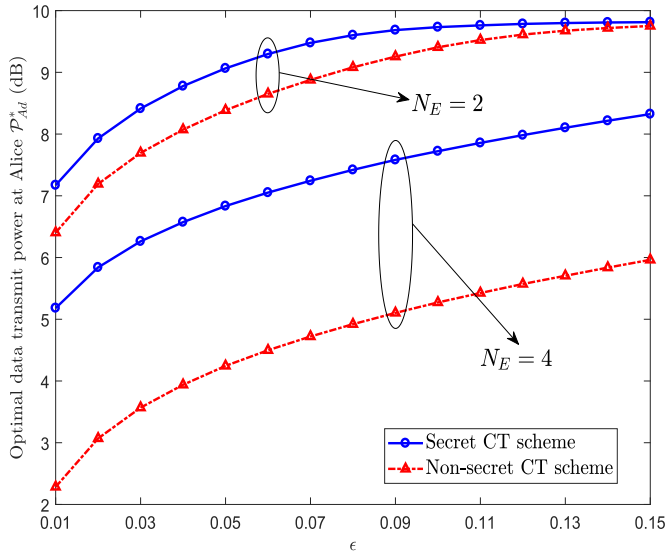


Fig. 6. Alice's optimal data transmit power $\mathcal{P}_{Ad}^*$ versus the secrecy constraint ϵ for different values of N_E, where $N_B = 8$, $R_B = 5$, $R_E = 3$, $T = 300$, $\sigma_s^2 = 1$, $\mathcal{E}_A/T = \mathcal{E}_B/T = 10$dB, and $\sigma_B^2 = 0$dB.

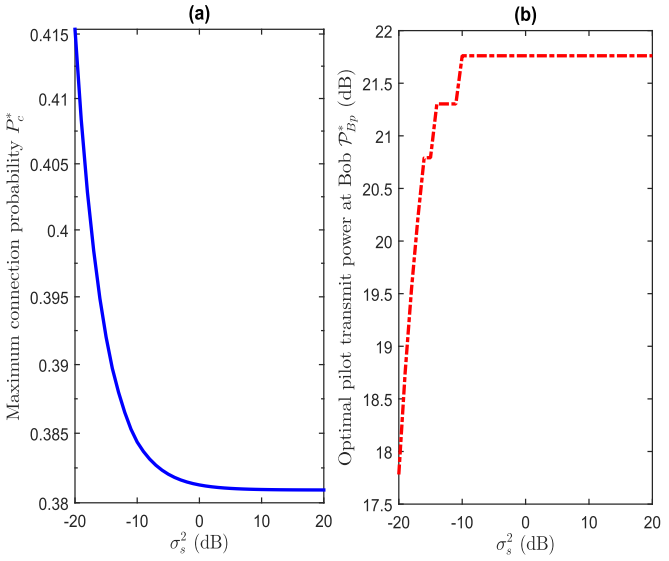


Fig. 8. The maximum connection probability and Bob's optimal pilot transmit power for the secret CT scheme versus the fading power of the self-interference channel σ_s^2, where $\epsilon = 0.10$, $R_B = 5$, $R_E = 3$, $T = 300$, $N_B = N_E = 2$, $\mathcal{E}_A/T = 15$dB, $\mathcal{E}_B/T = 10$dB, and $\sigma_B^2 = 0$dB.

respectively. We first observe that $\mathcal{P}_{Ad}^*$ increases as ϵ increases in Fig. 6, which demonstrates that more transmit power is allocated to data transmission as the secrecy constraint is relaxed. We also observe that $\mathcal{P}_{Ba}^*$ decreases as ϵ increases in Fig. 7, which demonstrates that as the secrecy constraint is relaxed less transmit power is allocated to AN at Bob. In Fig. 6, we further observe that more transmit power is allocated to data transmission at Alice as N_E decreases. In Fig. 7, we observe that less transmit power is allocated to AN at Bob as N_E decreases. As expected, in Fig. 6 we see that $\mathcal{P}_{Ad}^*$ for the secret CT scheme is higher than that for the non-secret CT scheme and in Fig. 7 we see that $\mathcal{P}_{Ba}^*$ for the secret CT scheme is lower than that for the non-secret CT scheme. This is due to the fact that Eve obtains less information regarding the jamming channel in the secret CT scheme.

In Fig. 8 we examine the impact of the fading power of the self-interference channel σ_s^2 (i.e., the variance of each entry in $\mathbf{H}_s$) on the secrecy performance and power allocation of the secret CT scheme. As expected, we observe that the maximum connection probability of the secret CT scheme decreases as σ_s^2 increases in Fig. 8(a). This is due to the fact that, according to (9), the channel estimation error of $\mathbf{H}_s$ increases as σ_s^2 increases, which means that the connection probability for the secret CT scheme is a monotonically decreasing function of σ_s^2. Noting the limited value range of the y-axis in Fig. 8(a),

we would like to highlight that the secrecy performance of the secret CT scheme is insensitive to σ_s^2, especially when $\sigma_s^2 > 0$dB. In Fig. 8(b), we observe that Bob's optimal pilot transmit power $\mathcal{P}_{Bp}^*$ first increases and then remains constant as σ_s^2 increases. The initial increase in $\mathcal{P}_{Bp}^*$ is due to the fact that as σ_s^2 increases the performance of the channel estimation decreases and thus more power should be allocated to the channel training to counteract this performance loss in the estimate of the self-interference channel. Again, noting the limited value range of the y-axis in Fig. 8(b), we can conclude that the power allocation of the secret CT scheme is insensitive to σ_s^2 as well.

In the secret CT scheme Bob can transmit AN during the training of the main channel to further enhance the physical layer security of the considered system, since the transmitted AN can increase the estimation error of the eavesdropper's channel. The role of AN in channel training in wiretap channels has been examined in the literature (e.g., [25], [35], and [36]). When Bob transmits AN during the training of the main channel, the estimation error of the main channel will depend on the realization of the self-interference channel and the estimation error of the eavesdropper's channel will depend on the realization of the jamming channel, which makes the closed-form performance analysis mathematically intractable. As such, we numerically approximate the performance of the updated secret CT scheme, in which the training of the self-interference channel is conducted before the training of the main channel and the transmission of AN by Bob is considered during the training of the main channel. As expected, we confirm that the transmission of AN by Bob during the training of the main channel improves the performance of the proposed scheme.

VI. Conclusion

In this work, we devised a new secret CT scheme based on the fact that in the wiretap channel with a full-duplex receiver, Bob knows exactly what he transmits. To study the performance of the proposed secret CT scheme, we derived closed-form expressions for the CP and SOP, based on which the power allocation between CT and data/AN transmission is optimized. Our examination demonstrates that when $N_E > 1$ the secret CT scheme significantly outperforms the non-secret CT scheme in terms of achieving a much higher CP subject to the same secrecy constraint, and when $N_E = 1$ they achieve the same secrecy performance. The secrecy performance improvement of the secret CT scheme relative to the non-secret CT scheme increases as N_E increases.

Appendix

Following (14), we can rewrite γ_B as

$$\gamma_B = \frac{\gamma_{B_n}}{\gamma_{B_s} + 1}, \tag{39}$$

where $\gamma_{B_n} = \mu_B \|\hat{\mathbf{h}}\|^2$, $\gamma_{B_s} = \gamma_{B_1} + \gamma_{B_2}$, $\gamma_{B_1} = \mu_B |\hat{\mathbf{h}}^\dagger \tilde{\mathbf{h}}|^2 / \|\hat{\mathbf{h}}\|^2$, and $\gamma_{B_2} = \mu_S \|\hat{\mathbf{h}}^\dagger \tilde{\mathbf{H}}_s\|^2 / \|\hat{\mathbf{h}}\|^2$. We note that here $\|\hat{\mathbf{h}}\|^2$ is a random variable since we are interested in the channel-independent CP, based on which the power allocation between channel estimation and data transmission can be determined. We note that γ_{B_n} and γ_{B_s} are independent due to the fact that $\hat{\mathbf{h}}$, $\tilde{\mathbf{h}}$, and $\tilde{\mathbf{H}}_s$ are independent of each other, and thus following (24) we have

$$\begin{aligned} P_c &= \Pr\left(\gamma_B \geq 2^{R_B} - 1\right) \\ &= \Pr\left(\gamma_{B_s} \leq \frac{\gamma_{B_n} - c}{c}\right) \\ &\overset{a}{=} \int_c^\infty F_{\gamma_{B_s}}\left(\frac{x-c}{c}\right) f_{\gamma_{B_n}}(x) dx \\ &= \int_0^\infty F_{\gamma_{B_s}}\left(\frac{x}{c}\right) f_{\gamma_{B_n}}(x+c) dx, \end{aligned} \tag{40}$$

where $\overset{a}{=}$ is achieved by noting $\gamma_{B_s} \geq 0$, $F_{\gamma_{B_s}}(\cdot)$ is the cdf of γ_{B_s}, $f_{\gamma_{B_n}}(\cdot)$ is the probability density function (pdf) of γ_{B_n}, and as defined in Theorem 1 we have $c = 2^{R_B} - 1$. We recall that the entries of $\hat{\mathbf{h}}$ are i.i.d and each of them follows the distribution $\mathcal{CN}(0, \sigma_{\hat{h}}^2)$, thus $f_{\gamma_{B_n}}(\cdot)$ is given by [43]

$$f_{\gamma_{B_n}}(x) = \frac{x^{N_B - 1}}{\Gamma(N_B) \mu_h^{N_B}} e^{-\frac{x}{\mu_h}}. \tag{41}$$

We next derive an expression for $F_{\gamma_{B_s}}(\cdot)$ when $\mu_t \neq \mu_H$, i.e., $F_{\gamma_{B_s}}(x) = \Pr(\gamma_{B_1} + \gamma_{B_2} \leq x)$. Since $\hat{\mathbf{h}}$ and $\tilde{\mathbf{h}}$ are independent, the cdf of γ_{B_1} is given by [43]

$$F_{\gamma_{B_1}}(x) = 1 - e^{\frac{x}{\mu_t}}. \tag{42}$$

Given that $\hat{\mathbf{h}}$ is independent of $\tilde{\mathbf{H}}_s$, the pdf of γ_{B_2} is given by [43]

$$f_{\gamma_{B_2}}(x) = \frac{x^{N_B - 1}}{\Gamma(N_B) \mu_H^{N_B}} e^{-\frac{x}{\mu_H}}. \tag{43}$$

We note that γ_{B_1} and γ_{B_2} are independent. As such, for $\mu_t \neq \mu_H$ we can derive $F_{\gamma_{B_s}}(x)$ as

$$\begin{aligned} &F_{\gamma_{B_s}}(x) \\ &= \int_0^x F_{\gamma_{B_1}}(x-y) f_{\gamma_{B_2}}(y) dy \\ &= \int_0^x \left[1 - e^{\frac{x-y}{\mu_t}}\right] f_{\gamma_{B_2}}(y) dy \\ &= \int_0^x f_{\gamma_{B_2}}(y) dy - \frac{e^{-\frac{x}{\mu_t}}}{\Gamma(N_B) \mu_H^{N_B}} \int_0^x y^{N_B - 1} e^{-\frac{(\mu_t - \mu_H) y}{\mu_t \mu_H}} dy \\ &\overset{b}{=} \frac{\gamma\left(N_B, \frac{x}{\mu_H}\right)}{\Gamma(N_B)} - \frac{\mu_t^{N_B} e^{-\frac{x}{\mu_t}} \gamma\left(N_B, \frac{(\mu_t - \mu_H) x}{\mu_t \mu_H}\right)}{\Gamma(N_B) (\mu_t - \mu_H)^{N_B}}, \end{aligned} \tag{44}$$

where $\overset{b}{=}$ is achieved with the aid of the following identity [46, eq. (3.351.1)]

$$\int_0^u x^n e^{-\mu x} dx = \mu^{-n-1} \gamma(n+1, \mu u), \tag{45}$$

and $\gamma(n, t)$ is the lower incomplete gamma function, which can be expanded for positive integer n as

$$\gamma(n, t) = (n-1)! \left[1 - e^{-t} \sum_{m=0}^{n-1} \frac{t^m}{m!}\right]. \tag{46}$$

Then, substituting (41) and (44) into (40), we have

$$\begin{aligned}P_c &= \frac{\mu_h^{-N_B}}{\Gamma(N_B)^2}\int_0^{\infty}(x+c)^{N_B-1}e^{-\frac{x+c}{\mu_h}}\gamma\left(N_B,\frac{x}{c\mu_H}\right)dx\\ &\quad-\frac{\mu_t^{N_B}\mu_h^{-N_B}}{\Gamma(N_B)^2}\int_0^{\infty}\frac{(x+c)^{N_B-1}}{e^{\frac{x+c}{\mu_h}+\frac{x}{c\mu_t}}}\gamma\left(N_B,\frac{(\mu_t-\mu_H)x}{c\mu_t\mu_H}\right)dx\\ &\overset{c}{=}\frac{e^{-\frac{c}{\mu_h}}\mu_h^{-N_B}}{\Gamma(N_B)^2}\sum_{i=0}^{N_B-1}\binom{N_B-1}{i}c^{N_B-i-1}\\ &\quad\times\left[\int_0^{\infty}x^ie^{-\frac{x}{\mu_h}}\gamma\left(N_B,\frac{x}{c\mu_H}\right)dx\right.\\ &\quad\left.-\frac{\mu_t^{N_B}}{(\mu_t-\mu_H)^{N_B}}\int_0^{\infty}\frac{x^i}{e^{\frac{c\mu_t+\mu_h}{c\mu_t\mu_h}x}}\gamma\left(N_B,\frac{(\mu_t-\mu_H)x}{c\mu_t\mu_H}\right)dx\right],\end{aligned}\tag{47}$$

where $\overset{c}{=}$ is obtained with the aid of the following identity [46, eq. (1.111)]

$$(a+x)^n=\sum_{i=0}^{n}\binom{n}{i}x^ia^{n-i}.\tag{48}$$

We then solve the integrals in (47) and obtain the results in (25) with the aid of the following identity [46, eq. (6.455.2)]

$$\begin{aligned}&\int_0^{\infty}\frac{x^{u-1}}{e^{\beta x}}\gamma(\upsilon,\alpha x)dx\\ &\quad=\frac{\alpha^{\upsilon}\Gamma(u+\upsilon)}{\upsilon(\alpha+\beta)^{u+\upsilon}}{}_2F_1\left(1,u+\upsilon;\upsilon+1;\frac{\alpha}{\alpha+\beta}\right),\\ &\qquad\qquad[\alpha+\beta>0,\beta>0,u+\upsilon>0].\end{aligned}\tag{49}$$

So far, we have proved (25) for $\mu_t\neq\mu_H$. We next prove (25) for the special case of $\mu_t=\mu_H$. Noting (42), (43), and that $\hat{\mathbf{h}}$ is independent of $\tilde{\mathbf{h}}$, for $\mu_t=\mu_H$ we have

$$\begin{aligned}F_{\gamma_{B_s}}(x)&=\frac{\gamma\left(N_B+1,\frac{x}{\mu_t}\right)}{\Gamma(N_B+1)}=\frac{\gamma\left(N_B+1,\frac{x}{\mu_H}\right)}{\Gamma(N_B+1)}\\ &\overset{d}{=}\frac{\gamma\left(N_B,\frac{x}{\mu_H}\right)}{\Gamma(N_B)}-\frac{\left(\frac{x}{\mu_H}\right)^{N_B}e^{-\frac{x}{\mu_H}}}{\Gamma(N_B+1)},\end{aligned}\tag{50}$$

where $\overset{d}{=}$ is achieved with the aid of $\gamma(N+1,x)=N\gamma(N,x)-x^Ne^{-x}$. Then, substituting (41) and (50) into (40), for $\mu_t=\mu_H$ we have

$$\begin{aligned}P_c &= \frac{\mu_h^{-N_B}}{\Gamma(N_B)^2}\int_0^{\infty}(x+c)^{N_B-1}e^{-\frac{x+c}{\mu_h}}\gamma\left(N_B,\frac{x}{c\mu_H}\right)dx\\ &\quad-\frac{\mu_h^{-N_B}(c\mu_H)^{-N_B}}{\Gamma(N_B)\Gamma(N_B+1)}\int_0^{\infty}x^{N_B}(x+c)^{N_B-1}e^{-\frac{x+c}{\mu_h}-\frac{x}{c\mu_H}}dx\\ &=\frac{e^{-\frac{c}{\mu_h}}\mu_h^{-N_B}}{\Gamma(N_B)^2}\sum_{i=0}^{N_B-1}\binom{N_B-1}{i}c^{N_B-i-1}\\ &\quad\times\left[\int_0^{\infty}x^ie^{-\frac{x}{\mu_h}}\gamma\left(N_B,\frac{x}{c\mu_H}\right)dx\right.\\ &\quad\left.-\frac{(c\mu_H)^{-N_B}}{N_B}\int_0^{\infty}x^{N_B+i}e^{-\frac{c\mu_H+\mu_h}{c\mu_H\mu_h}x}dx\right]\\ &\overset{e}{=}\frac{c^{N_B}e^{-\frac{c}{\mu_h}}}{\Gamma(N_B+1)\Gamma(N_B)(c\mu_H+\mu_h)^{N_B}}\\ &\quad\times\sum_{i=0}^{N_B-1}\frac{\binom{N_B-1}{i}\Gamma(N_B+i+1)}{(c\mu_H+\mu_h)^{i+1}(\mu_H\mu_h)^{-i-1}}\\ &\quad\times\left[{}_2F_1\left(1,N_B+i+1;N_B+1;\frac{\mu_h}{c\mu_H+\mu_h}\right)-1\right],\end{aligned}\tag{51}$$

where $\overset{e}{=}$ is achieved with the aid of (49) and the following identity [46, eq. (3.351.3)]

$$\int_0^{\infty}x^ne^{-ux}dx=n!u^{-n-1},\quad u>0.\tag{52}$$

We note that for $\mu_t=\mu_H$ in (25) we have

$${}_2F_1\left(1,N_B+i+1;N_B+1;\frac{\mu_h(\mu_t-\mu_H)}{\mu_t(c\mu_H+\mu_h)}\right)=1.\tag{53}$$

As such, comparing (51) with (25) we know that (25) is also valid for the special case of $\mu_t=\mu_H$. Therefore, following (47) and (51) the proof of Theorem 1 is completed.

REFERENCES

[1] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] X. Zhou, L. Song, and Y. Zhang, *Physical Layer Security in Wireless Communications*. Boca Raton, FL, USA: CRC Press, 2013.

[3] 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, 3rd Quart. 2014.

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

[5] K. Ren, H. Su, and Q. Wang, "Secret key generation exploiting channel characteristics in wireless communications," *IEEE Wireless Commun.*, vol. 18, no. 4, pp. 6–12, Aug. 2011.

[6] C. E. Shannon, "Communication theory of secrecy systems," *Bell Labs Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

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

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

[9] S. A. A. Fakoorian and A. L. Swindlehurst, "Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5013–5022, Oct. 2011.

[10] H.-M. Wang, T. Zheng, and X.-G. Xia, "Secure MISO wiretap channels with multiantenna passive eavesdropper: Artificial noise vs. Artificial fast fading," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 94–106, Jan. 2015.

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

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

[13] S. Yan, X. Zhou, N. Yang, B. He, and T. D. Abhayapala, "Artificial-noise-aided secure transmission in wiretap channels with transmitter-side correlation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8286–8297, Dec. 2016.

[14] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, "Improving physical layer secrecy using full-duplex jamming receivers," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4962–4974, Oct. 2013.

[15] L. Li, Z. Chen, D. Zhang, and J. Fang, "A full-duplex Bob in the MIMO Gaussian wiretap channel: Scheme and performance," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 107–111, Jan. 2016.

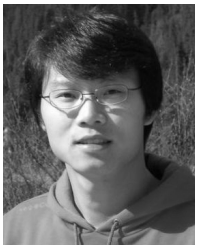
[16] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[17] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM*, Aug. 2013, pp. 375–386.

[18] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[19] W. Li, M. Ghogho, B. Chen, and C. Xiong, "Secure communication via sending artificial noise by the receiver: Outage secrecy capacity/region analysis," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1628–1631, Oct. 2012.
[20] Y. Zhou, Z. Z. Xiang, Y. Zhu, and Z. Xue, "Application of full-duplex wireless technique into secure MIMO communication: Achievable secrecy rate based optimization," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 804–808, Jul. 2014.
[21] S. Yan, N. Yang, R. Malaney, and J. Yuan, "Full-duplex wiretap channels: Security enhancement via antenna switching," in *Proc. IEEE GlobeCOM TCPLS Workshop*, Dec. 2014, pp. 1412–1417.
[22] F. Zhu, F. Gao, T. Zhang, K. Sun, and M. Yao, "Physical-layer security for full duplex communications with self-interference mitigation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 329–340, Jan. 2016.
[23] X. Tang, P. Ren, Y. Wang, and Z. Han, "Combating full-duplex active eavesdropper: A hierarchical game perspective," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1379–1395, Mar. 2017.
[24] J. Hu, Y. Cai, N. Yang, X. Zhou, and W. Yang, "Artificial-noise-aided secure transmission scheme with limited training and feedback overhead," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 193–205, Jan. 2017.
[25] T.-Y. Liu, S.-C. Lin, and Y.-W. P. Hong, "On the role of artificial noise in training and data transmission for secret communications," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 3, pp. 516–531, Mar. 2017.
[26] S. Yan, X. Zhou, N. Yang, T. D. Abhayapala, and A. L. Swindlehurst, "Channel training design in full-duplex wiretap channels to enhance physical layer security," in *Proc. IEEE ICC*, May 2017, pp. 1–6.
[27] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4381–4393, Dec. 2011.
[28] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Physical layer network security in the full-duplex relay system," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 574–583, Mar. 2015.
[29] S. Dang, G. Chen, and J. P. Coon, "Outage performance analysis of full-duplex relay-assisted device-to-device systems in uplink cellular networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4506–4510, May 2017.
[30] 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.
[31] G. Zhang, J. Xu, Q. Wu, M. Cui, X. Li, and F. Lin, "Wireless powered cooperative jamming for secure OFDM system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1331–1346, Feb. 2018.
[32] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[33] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[34] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.
[35] T.-H. Chang, W.-C. Chiang, Y.-W. P. Hong, and C.-Y. Chi, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6223–6237, Dec. 2010.
[36] C.-W. Huang, X. Zhou, T.-H. Chang, and Y.-W. P. Hong, "Two-way training for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2724–2738, May 2013.
[37] J. Xu and R. Zhang, "Energy beamforming with one-bit feedback," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5370–5381, Oct. 2014.
[38] V. S. Annapureddy, D. V. Marathe, T. R. Ramya, and S. Bhashyam, "Outage probability of multiple-input single-output (MISO) systems with delayed feedback," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 319–326, Feb. 2009.
[39] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.
[40] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.
[41] S. Yan, N. Yang, G. Geraci, R. Malaney, and J. Yuan, "Optimization of code rates in SISOME wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6377–6388, Nov. 2015.
[42] X. Zhou, M. R. McKay, B. Maham, and A. Hjørungnes, "Rethinking the secrecy outage formulation: A secure transmission design perspective," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 302–304, Mar. 2011.
[43] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[44] L. Lai, H. E. Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.
[45] B. He and X. Zhou, "Secure on-off transmission design with channel estimation errors," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 1923–1936, Dec. 2013.
[46] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 7th ed. San Diego, CA, USA: Academic, 2007.

Shihao Yan (S'11–M'15) received the B.S. degree in communication engineering and the M.S. degree in communication and information systems from Shandong University, Jinan, China, in 2009 and 2012, respectively, and the Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, Australia, in 2015. From 2015 to 2017, he was a Post-Doctoral Research Fellow with the Research School of Engineering, Australia National University, Canberra, Australia. He is currently a University Research Fellow with the School of Engineering, Macquarie University, Sydney, Australia. His current research interests are in the areas of wireless communications and statistical signal processing, including physical layer security, covert communications, and location spoofing detection.



Xiangyun Zhou (SM'17) received the Ph.D. degree from Australian National University (ANU) in 2010. He is currently a Senior Lecturer with ANU. His research interests are in the fields of communication theory and wireless networks. He was a recipient of the Best Paper Award at ICC'11 and the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2016. He was named the Best Young Researcher in the Asia-Pacific Region in 2017 by IEEE ComSoc Asia-Pacific Board. He also served as symposium/track and workshop co-chairs for major IEEE conferences. He was the Chair of the ACT Chapter of the IEEE Communications Society and Signal Processing Society from 2013 to 2014. He has been serving as an Editor for various IEEE journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE COMMUNICATIONS LETTERS. He served as a Guest Editor for the *IEEE Communications Magazine*'s feature topic on wireless physical layer security in 2015.

Nan Yang (S'09–M'11) received the B.S. degree in electronics from China Agricultural University in 2005, and the M.S. and Ph.D. degrees in electronic engineering from the Beijing Institute of Technology in 2007 and 2011, respectively. He was a Post-Doctoral Research Fellow with the Commonwealth Scientific and Industrial Research Organization from 2010 to 2012 and the University of New South Wales from 2012 to 2014. He has been with the Research School of Engineering, Australian National University, since 2014, where he is currently a Future Engineering Research Leadership Fellow and a Senior Lecturer. His general research interests include massive multi-antenna systems, millimeter-wave and terahertz communications, ultra-reliable low latency communications, cyber-physical security, and molecular communications. He received the Top Editor Award from the *Transactions on Emerging Telecommunications Technologies* in 2017, the Exemplary Reviewer Award from the IEEE TRANSACTIONS ON COMMUNICATIONS in 2016 and 2015, the Top Reviewer Award from the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2015, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award and the Exemplary Reviewer Award from the IEEE WIRELESS COMMUNICATIONS LETTERS in 2014, and the Exemplary Reviewer Award from the IEEE COMMUNICATIONS LETTERS in 2013 and 2012. He was also a co-recipient of the Best Paper Awards from IEEE GLOBECOM 2016 and IEEE VTC 2013-Spring. He is currently serving in the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and *Transactions on Emerging Telecommunications Technologies*.

Thushara D. Abhayapala (M'00–SM'08) received the B.E. degree (Hons.) in engineering and the Ph.D. degree in telecommunications engineering from the Australian National University (ANU), Canberra, in 1994 and 1999, respectively. He was the Director of the Research School of Engineering, ANU, from 2010 to 2014, and also the Leader of the Wireless Signal Processing Program with the National ICT Australia from 2005 to 2007. He is currently the Deputy Dean of the College of Engineering and Computer Science, ANU. He has supervised over 30 Ph.D. students and co-authored over 200 peer reviewed papers. His research interests are in the areas of spatial audio and acoustic signal processing, multi-channel signal processing, and spatial aspects of wireless communications. He is a member of the Audio and Acoustic Signal Processing Technical Committee of the IEEE Signal Processing Society from 2011 to 2016. He is an Associate Editor of IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING. He is a Fellow of Engineers Australia.

A. Lee Swindlehurst (F'04) received the B.S. and M.S. degrees in electrical engineering from Brigham Young University (BYU), in 1985 and 1986, respectively, and the Ph.D. degree in electrical engineering from Stanford University in 1991. He was with the Department of Electrical and Computer Engineering, BYU, from 1990 to 2007, where he served as the Department Chair from 2003 to 2006. From 1996 to 1997, he held a visiting scholar position with Uppsala University and the Royal Institute of Technology, Sweden. From 2006 to 2007, he was on leave, while he was a Vice President of Research for ArrayComm LLC, San Jose, CA, USA. Since 2007, he has been a Professor with the Electrical Engineering and Computer Science Department, University of California at Irvine, Irvine, where he served as an Associate Dean for Research and Graduate Studies with the Samueli School of Engineering from 2013 to 2016. From 2014 to 2017, he was also a Hans Fischer Senior Fellow with the Institute for Advanced Studies, Technical University of Munich. His research focuses on array signal processing for radar, wireless communications, and biomedical applications, and he has over 300 publications in these areas. He was a recipient of the 2000 IEEE W. R. G. Baker Prize Paper Award, the 2006 IEEE Communications Society Stephen O. Rice Prize in the Field of Communication Theory, the 2006 and 2010 IEEE Signal Processing Society's Best Paper Awards, and the 2017 IEEE Signal Processing Society Donald G. Fink Overview Paper Award. He was the inaugural Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.