# A Broadcast Approach To Secret Key Generation Over Slow Fading Channels

Xiaojun Tang, Ruoheng Liu, Predrag Spasojević and H. Vincent Poor

***Abstract*—A secret-key generation scheme based on a layered broadcasting strategy is introduced for slow-fading channels. In the model considered, Alice wants to share a key with Bob while keeping the key secret from Eve, who is a passive eavesdropper. Both Alice-Bob and Alice-Eve channels are assumed to undergo slow fading, and perfect channel state information (CSI) is assumed to be known only at the receivers during the transmission. In each fading slot, Alice broadcasts a continuum of coded layers and, hence, allows Bob to decode at the rate corresponding to the fading state (unknown to Alice). The index of a reliably decoded layer is sent back from Bob to Alice via a public and error-free channel and used to generate a common secret key. In this paper, the achievable secrecy key rate is first derived for a given power distribution over coded layers. The optimal power distribution is then characterized. It is shown that layered broadcast coding can increase the secrecy key rate significantly compared to single-level coding.**

***Index Terms*—Secret-key agreement, wiretap channel, layered broadcast coding, superposition coding, feedback, interference, fading channel**

## I. Introduction

Wireless secrecy has attracted considerable research interest due to the concern that wireless communication is highly vulnerable to security attacks, particularly eavesdropping attacks. Much recent research was motivated by Wyner's wire-tap channel model [1], in which the transmission between two legitimate users (Alice and Bob) is eavesdropped upon by Eve via a degraded channel. In this model, to characterize the leakage of information to the eavesdropper, equivocation rate is used to denote the level of ignorance of the eavesdropper with respect to the confidential messages. Perfect secrecy requires that the equivocation rate is asymptotically equal to the message rate, and the maximal achievable rate with perfect secrecy is called the secrecy capacity. Wyner showed that secret communication is possible without a secret-key shared by legitimate users. Later, Csiszár and Körner generalized Wyner's model to consider general broadcast channels in [2]. The Gaussian wire-tap channel was considered in [3]. Recent research has addressed the information-theoretic secrecy for multi-user channel models [4]–[9]. We refer the reader to [10] for a recent survey of the research progress in this area.

Interestingly, the wireless medium provides its own endowments that facilitate defending against eavesdropping. One such endowment is fading [11]. The effect of fading on secret transmission has been studied in [12]–[14]. In these works, assuming that all communicating parties have perfect channel state information (CSI), the ergodic secrecy capacity has been derived. The scenario in which Alice has no CSI about Eve's channel (but knows the channel statistics) has also been studied in [12]. The throughput of several secure hybrid automatic repeat request (ARQ) protocols has been analyzed in [15]. In this work, Alice is not assumed to have prior CSI (except channel statistics), but can receive a 1-bit ARQ feedback per channel coherence interval from Bob reliably.

Arguably, the most useful application of (keyless) secret message transmission is secret-key generation. For instance, a key can be sent from Alice to Bob as a secret message (which is selected by Alice in advance). More generally, as considered here, the key can be established after a communication session completes. This relaxation in the protocol can lead to a higher key rate. The secret-key generation problem in [16] and [17] assumes an interactive, authenticated public channel with unlimited capacity. In [17], the "channel model with wiretapper" (CW) is similar to the wiretap channel model, while in the "source model with wiretapper" (SW), Alice and Bob exploit correlated source observations to generate the key. Both SW and CW models have been subsequently extended to multiple terminals [18]–[20] and to non-authenticated public channels [21]–[23]. Secret-key generation using both correlated sources and channels has been considered more recently in [24] and [25].

In this paper, we consider a key-generation problem in which Alice wants to share a key with Bob while keeping it secret from Eve. The Alice-Bob and Alice-Eve channels (forward channels) undergo slow fading, and CSI is known only at the receivers. Furthermore, we assume a public and error-free feedback channel. The key generation scheme under consideration consists of a communication and a key-

This research was supported by the National Science Foundation under Grants CNS-09-05398, CCF-07-28208 and CCF-0729142, and in part by the Air Force Office of Scientific Research under Grant FA9550-08-1-0480. The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Seoul, Korea, June 24 - 29, 2009.

Xiaojun Tang is with AT&T Labs, San Ramon, CA 94583 USA (e-mail: xiaojun.tang@att.com). Ruoheng Liu is with Alcatel-Lucent, Murray Hill, NJ 07974 USA. (email: ruoheng.liu@alcatel-lucent.com). Predrag Spasojević is with the Wireless Information Network Laboratory (WINLAB), Department of Electrical and Computer Engineering, Rutgers University, North Brunswick, NJ 08902, USA (e-mail: spasojev@winlab.rutgers.edu). H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu).

generation phase. In the communication phase, via the forward channel, Alice sends to Bob coded sequences, which are observed at Bob and Eve after independent distortions due to power attenuation and noise. Subsequently, Alice and Bob agree on the same secret-key in the key-generation phase. The problem setting resembles an SW model but differs in that the shared "correlated sources" are coded sequences (from a public codebook and distorted by the channel). We assume that the feedback channel from Bob to Alice is very limited. For each block transmission from Alice to Bob, Bob is required to send back one or more bits to Alice, where the one-bit feedback corresponds to an ARQ ACK/NACK scheme. An example application is where Alice sends a video clip to Bob, which is a non-secret transmission. Bob responds with a few bits and thus enables agreeing on a secret-key, which can then be used in key-based cryptographic protocols.

The communication phase is based on layered broadcast coding, which effectively adapts the decoded rate at Bob to the actual channel state without requiring CSI to be available at Alice. The transmission takes place over several time slots. In each time slot, Alice transmits a continuum of layers. Depending on the realization of the channel state, Bob decodes a subset of layers reliably. The index of the highest reliably decoded layer at Bob is sent back to Alice, and used in the key-generation phase that follows Wyner's secrecy binning scheme [1]. For a given power distribution over coded layers, we derive the achievable secrecy key rate, which permits a simple interpretation as the average reward collected from all possible channel realizations. Furthermore, we characterize the optimal power distribution over coded layers to maximize the achievable secrecy key rate under the broadcast approach.

Layered broadcast coding creates *artificial noise* so that the undecodable layers at Bob play the role of self-interference. We show that, by properly choosing the coding rate for each layer, it is ensured that Eve cannot benefit from the layered coding structure and is forced to treat the layers undecodable at Bob as interference. Secret communications with interference was studied in [26] and [27] in a more general (but non-fading) setting. Layered broadcast coding for a slow-fading single-input single-output (SISO) channel model was originally introduced by Shamai in [28] and discussed in greater details in [29]. The results in this paper are consistent with [28] and [29] when the additional secrecy key generation requirement phase is not considered. In a closely related work, a similar ARQ-based secret-key generation scheme employing single-level coding was studied in [30]. This scheme can be viewed as a special case of the proposed layered-coding based scheme as all power is allocated to a single coded layer. We show that layered broadcast coding can increase the secrecy key rate significantly compared to single-level coding.

The remainder of the paper is organized as follows. Section II describes the system model. Section III states the broadcast approach for key generation. Section IV gives the achievable secrecy rate for a given power distribution. Section V characterizes the optimal power distribution. A numerical example involving a Rayleigh fading channel is given in Section VI. Conclusions are given in Section VII.

## II. System Model

As depicted in Fig. 1, we consider a three-terminal model, in which Alice and Bob want to share a secret key in the presence of Eve, who is a passive eavesdropper. That is, Eve is interested in stealing the key but does not attempt to interfere with the key generation processes.

### A. Channel Model

The Alice-Bob and Alice-Eve channels (forward channels) undergo block fading, in which the channel gains are constant within a block while varying independently from block to block [11]. We assume that each block is associated with a time slot of duration $T$ and bandwidth $B$; that is, $N = \lfloor 2BT \rfloor$ real symbols can be sent in each slot. We also assume that the number of channel uses within each slot (i.e., $N$) is large enough to allow for invoking random coding arguments.

Let us assume that the transmissions in the forward channels take place over $M$ time slots. In a time slot indexed by $m \in [1, \ldots, M]$, Alice sends $\mathbf{X}_m$, which is a vector of $N$ real symbols. Bob receives $\mathbf{Y}_{1m}$ through the channel gain $h_{1m}$ and Eve receives $\mathbf{Y}_{2m}$ through the channel gain $h_{2m}$. A discrete time baseband-equivalent block-fading channel model can be expressed as

$$\mathbf{Y}_{tm} = \sqrt{h_{tm}}\mathbf{X}_m + \mathbf{Z}_{tm} \tag{1}$$

for $t = 1, 2$, where $\{\mathbf{Z}_{tm}\}$ are sequences of independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian $\mathcal{N}(0, 1)$ random variables. We denote by $h_{1m}$ and $h_{2m}$ the states of the Alice-Bob and Alice-Eve channels, respectively, in time slot $m$. Without loss of generality, we drop the index $m$ and denote random channel realizations by $h_t$. We assume that $h_t$ is a real random variable with a probability density function (PDF) $f_t$ and a cumulative distribution function (CDF) $F_t$, for each $t = 1, 2$. We also let $\mathbf{h}_1 = [h_{1,1}, \ldots, h_{1,M}]$ and $\mathbf{h}_2 = [h_{2,1}, \ldots, h_{2,M}]$ denote the power gain vectors for the Alice-Bob and Alice-Eve channels, respectively. We assume that Bob and Eve know their own channel gains perfectly; Alice does not know the CSI before its transmission, except for the channel statistics.

In addition, we assume a short term power constraint (excluding power variation across time slots) such that the average power of the signal $\mathbf{X}_m$ per slot satisfies the constraint

$$\frac{1}{N} E[\|\mathbf{X}_m\|^2] \leq P \tag{2}$$

for all $m = 1, \ldots, M$.

Finally, we assume that there exists an error-free feedback channel from Bob to Alice, through which Bob can feed

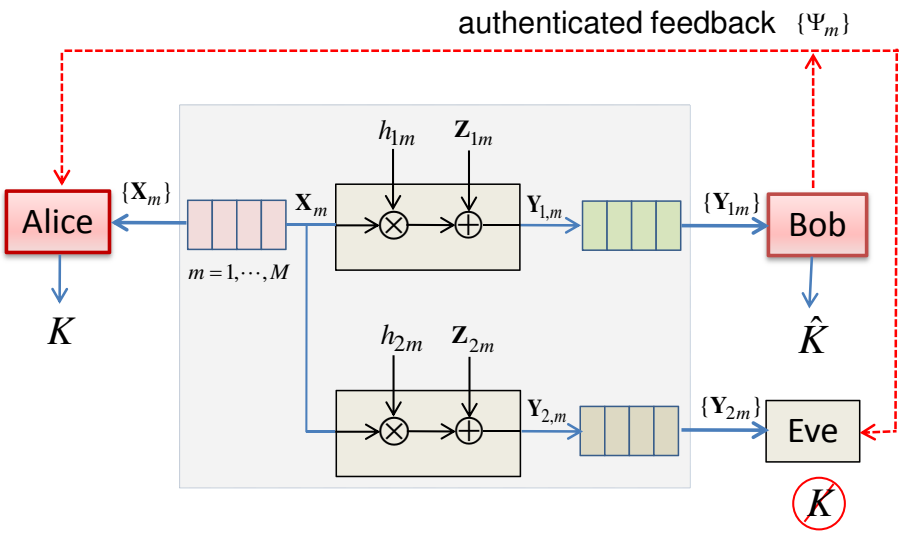


Fig. 1. Alice and Bob want to agree on a key ($K=\hat{K}$), while keeping the key secret from Eve ($H(K|\mathbf{Y}_2,\mathbf{h}_2,\mathbf{\Psi})/n\to 0$).

back $\Psi_m$ for time slot $m$, where $\Psi_m$ is a deterministic function of $\mathbf{Y}_{1m}$ and $h_{1,m}$. The feedback channel is assumed to be public, and therefore $\Psi_m$ is received by both Alice and Eve without any error.

### B. Secret Key Generation Protocol

The secret key generation protocol consists of two phases: a communication phase and a key-generation phase.

*1) Communication Phase:* We assume that the transmission during the communication phase takes place over $M$ time slots. That is, Alice sends a sequence of signals $\mathbf{X}=(\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_M)$ to the channel. Accordingly, Bob receives from his channel a sequence of signals denoted by $\mathbf{Y}_1=(\mathbf{Y}_{1,1},\mathbf{Y}_{1,2},\ldots,\mathbf{Y}_{1,M})$ and Eve receives $\mathbf{Y}_2=(\mathbf{Y}_{2,1},\mathbf{Y}_{2,2},\ldots,\mathbf{Y}_{2,M})$ from her channel. We let $n=MN$ denote the number of symbols sent by Alice in the communication phase.

After the transmission, Bob uses the feedback channel to send $\mathbf{\Psi}=(\Psi_1,\ldots,\Psi_M)$, which is received by both Alice and Eve since the feedback channel is public and error-free.

*2) Key-Generation Phase:* The communication phase is followed by a key-generation phase, in which both Alice and Bob generate the key based on the forward and backward signals. A general key-generation phase can be described as in the following.

Let $\mathcal{K}=\{1,2,\ldots,2^{nR_s}\}$, where $R_s$ represents the secrecy key rate. Alice generates a secret key $k\in\mathcal{K}$ by using a decoding function $K$, i.e.,

$$k=K\left(\mathbf{X},\mathbf{\Psi}\right). \tag{3}$$

Bob generates the secret key $\hat{k}\in\mathcal{K}$ by using a decoding function $\hat{K}$, i.e.,

$$\hat{k}=\hat{K}\left(\mathbf{Y}_1,\mathbf{h}_1,\mathbf{\Psi}\right)=\hat{K}\left(\mathbf{Y}_1,\mathbf{h}_1\right), \tag{4}$$

where the second equality holds since we assume that $\mathbf{\Psi}$ is a deterministic function of $\mathbf{Y}_1$ and $\mathbf{h}_1$.

The secrecy level at Eve is measured by the equivocation rate $R_e$ defined as the entropy rate of the key $K$ conditioned upon the observations at Eve, i.e.,

$$R_e\triangleq\frac{1}{n}H(K|\mathbf{Y}_2,\mathbf{h}_2,\mathbf{\Psi}). \tag{5}$$

**Definition 1.** A secrecy key rate $R_s$ is achievable if the conditions

$$\Pr\left(K=\hat{K}\right)\geq 1-\epsilon, \tag{6}$$

$$\text{and}\qquad R_e\geq R_s-\epsilon, \tag{7}$$

are satisfied for any $\epsilon>0$ as the number of channel uses $n\to\infty$.

## III. A Layered Broadcast Approach To Key Generation

In this section, we introduce a broadcast approach for secret-key generation, in which Gaussian layered broadcast coding is used for the communication phase, and random secrecy binning is used for the key generation phase.

Before presenting the scheme, we briefly introduce Gaussian layered broadcast coding. Finite-level layered broadcast coding (superposition coding) was introduced by Cover in [31] for general broadcast channels. In [28], Shamai studied a Gaussian fading channel with no CSI at the transmitter and considered the limiting case when there is a continuum of coded layers. In this section, we first take a look at a fading wiretap channel with a finite number of fading states, for which finite level layered broadcast coding is applicable. The channel will be used to derive the result for the limiting case of continuous fading, which is the focus of this paper.

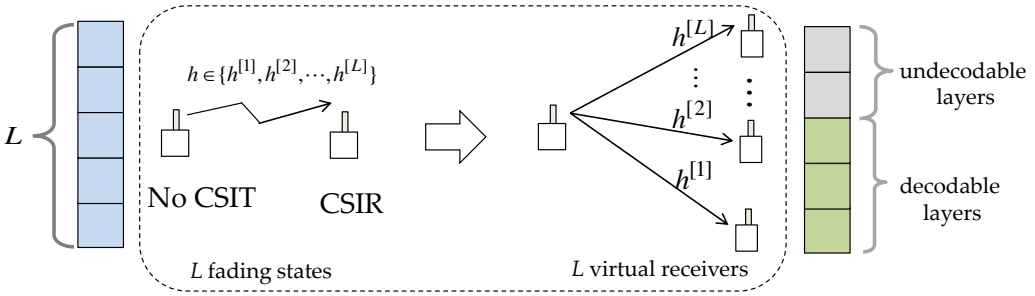


Fig. 2. A point-to-point fading channel with $L$ possible fading states is viewed as a broadcast channel with $L$ virtual receivers each corresponding to a fading state.

### *A. Finite-Level Layered Broadcast Coding for L-State Fading Channel*

Let us first consider a type of channel called "the $L$-state fading wiretap channel," in which there are $L$ different fading states possibly observed on the Alice-Bob or Alice-Eve channel.

**Definition 2.** In an $L$-state fading wiretap channel, at any time slot, the realization of the power gain of the Alice-Bob or Alice-Eve channel takes one value from $\{h^{[1]}, h^{[2]}, \ldots, h^{[L]}\}$ independently and randomly, and is characterized by probability function $\Pr\{h_1 = h^{[l_1]}, h_2 = h^{[l_2]}\}$. Without loss of generality, we assume that $\{h^{[l]}\}$ are ordered in ascending order.

Here, let us focus on the Alice-Bob channel. As shown in Fig. 2, in a layered broadcast coding scheme, the point-to-point fading channel is viewed as a broadcast channel with $L$ virtual receivers each corresponding to a fading state. By applying the superposition coding in [31], the encoding and decoding procedures can be described as follows.

During the encoding, we assume that $L$ layers are used. That is, the transmitted codeword is a superposition of $L$ codewords, i.e., $\sum_{l=1}^{L} \mathbf{X}^{[l]}$, where $\mathbf{X}^{[l]}$ is a codeword from a Gaussian codebook $\mathcal{C}^{[l]}$ with a rate $r^{[l]}$ and a constant power $p^{[l]}$, $l = 1, \ldots, L$. For a given power allocation $\{p^{[l]}\}$, the rate of the $l$-th layer is given by [1]

$$r^{[l]} = \log\left(1 + \frac{h^{[l]} p^{[l]}}{1 + h^{[l]} \sum_{i=l+1}^{L} p^{[i]}}\right), \tag{8}$$

and the total power satisfies $\sum_{l=1}^{L} p^{[l]} = P$.

During the decoding, for a given fading realization $h^{[l]}$, the receiver can successfully decode the first $l$ layers by using the successive decoding strategy [31]. i.e., the codewords $\{\mathbf{X}^{[1]}, \ldots, \mathbf{X}^{[l]}\}$ can be decoded reliably, while the codewords $\{\mathbf{X}^{[l+1]}, \ldots, \mathbf{X}^{[L]}\}$ are undecodable. More specifically, in the decoding process, the receiver first decodes $\mathbf{X}^{[1]}$ by treating the remaining codewords ($\{\mathbf{X}^{[i]}, i > 1\}$) as interference. After decoding $\mathbf{X}^{[1]}$, the receiver will subtract $\mathbf{X}^{[1]}$ and then decode $\mathbf{X}^{[2]}$ by treating the remaining codewords ($\{\mathbf{X}^{[i]}, i > 2\}$) as interference. This process repeats until the $l$-th layer $\mathbf{X}^{[l]}$ is decoded reliably by treating the remaining codewords ($\{\mathbf{X}^{[i]}, i > l\}$) as interference. As shown in (8), $\sum_{i=l+1}^{L} p^{[i]}$ is the total power of coded layers treated as inference during the decoding of the $l$-th layer. Note that this predetermined ordering can be achieved because of the degraded nature of Gaussian single-input single-output (SISO) channels.

[1]All logarithms are to the natural base, and thus rates are in terms of nats per second per Hertz.

### *B. Layered Broadcast Coding for Gaussian Fading Channels*

In general, $L$ depends on the cardinality of the random channel variable. For a Gaussian fading channel, a continuum of code layers ($L \to \infty$) is required for achieving the best performance. When a continuum of layers is used, the transmitter sends an infinite number of layers of coded information. Each layer conveys a fractional rate, denoted by $dR$, whose value depends on the index of the layer. We refer to $s$, the realization of the fading power, as a continuous index. For a given transmit power distribution $\rho(s)$ over coded layer $s$, $\rho(s)ds$ is the transmit power used by layer $s$. Any layer indexed by $u$ satisfying $u > s$ is undecodable and functions as additional interference. The total power of undecodable layers (for a realization of fading power $s$) is denoted by $I(s)$ and is expressed by

$$I(s) = \int_{s}^{\infty} \rho(u)du. \tag{9}$$

The incremental differential rate of layer $s$ is given by

$$dR(s) = \log\left(1 + \frac{s\rho(s)ds}{1 + sI(s)}\right) = \frac{s\rho(s)ds}{1 + sI(s)}, \tag{10}$$

where the second equality in (10) is due to the fact that $\lim_{x\to 0} \log(1+x) = x$ for any $x \geq 0$. The total power over all layers is constrained by

$$I(0) = \int_{0}^{\infty} \rho(u)du = P. \tag{11}$$

Given a realization of the fading power (or layer index) $s$, the decodable rate at the receiver is

$$R(s) = \int_0^s \frac{u\rho(u)du}{1+uI(u)}. \tag{12}$$

Hence, for a given CDF of the random fading power $s$ denoted by $F(s)$, the average decodable rate at the receiver is

$$R = \int_0^\infty \int_0^s \frac{u\rho(u)du}{1+uI(u)} dF(s). \tag{13}$$

### *C. Secret-Key Generation Based on Layered Broadcast Coding*

In this section, we discuss key generation based on Gaussian layered broadcast coding. We outline the scheme for the continuous case when $L \to \infty$, which is the focus of this paper. For an $L$-state fading wiretap channel when $L$ is finite, the corresponding scheme is discussed in Appendix A.

*1) Codebook Construction:* We need two types of codebooks used for the communication and key-generation phases, respectively.

The codebook used for the communication phase consists of a continuum of coded layers represented by $\{\mathcal{C}^{[s]}(2^{NdR(s)}, N)\}$, where $N$ is the codeword length and $dR(s)$ is the (incremental differential) rate at layer $s$. The (sub-)codebook for each layer is generated randomly and independently. That is, for any codebook $\mathcal{C}^{[s]}(2^{NdR(s)}, N)$, we generate $2^{NdR(s)}$ codewords $\mathbf{X}^{[s]}(w)$, where $w = 1, 2, \ldots, 2^{NdR(s)}$, by choosing the $N2^{NdR(s)}$ Gaussian symbols (with power $\rho(s)ds$) independently at random.

The codebook used for the key generation phase is based on Wyner's secrecy coding [1], [12]. As shown in Fig. 3, we use

$$R = \int_0^\infty \int_0^{h_1} \frac{s\rho(s)ds}{1+sI(s)} dF_1(h_1) \tag{14}$$

to represent the average decodable rate at Bob. We first generate all binary sequences of length $n(R-\epsilon)$, denoted by $\mathcal{B}$, where $n = MN$. The sequences $\mathcal{B}$ are then randomly and uniformly grouped into $K = 2^{nR_s}$ bins each with $n(R - R_s - \epsilon)$ sequences, where $R_s$ is the achievable secrecy rate given later. We denote by $\mathbf{v}(k, j)$ the $j$-th codeword in the $k$-th bin, where $1 \le k \le K$ and $1 \le j \le J = 2^{n(R-R_s-\epsilon)}$. Each secret key $k \in \{1, \ldots, K\}$ is then randomly assigned to a bin, denoted by $\mathcal{B}(k) = \{\mathbf{v}(k, j), j = 1, \ldots, J\}$.

*2) Communication Phase:* The communication takes places over $M$ time slots. In time slot $m \in [1, \ldots, M]$, Alice first randomly selects a message $W_m^{[s]} \in \{1, \ldots, 2^{NdR(s)}\}$ for coded layer $s$, independent of the message chosen for other layers. For convenience, we use $W_m$ to represent the total message sent in time slot $m$ (through all layers), i.e., $W_m = \times_s W_m^{[s]}$. Then, Alice sends a superposition of all layers to the channel.

Bob receives $\mathbf{Y}_{1m}$ and tries to decode all his decodable layers, which depends on his channel state $h_{1m}$. For convenience, we use $W_m^{[\mathcal{D}_1]}$ to denote the set of layers reliably decoded by Bob, and $W_m^{[\mathcal{U}_1]}$ to denote the set of layers undecodable to Bob in time slot $m$.[2] After decoding, Bob sends back the index of the highest decodable layer to Alice via the feedback channel, so that both Alice and Bob get to know $W_m$. This completes the transmission in time slot $m$. The communication phase ends when all $M$ (independent) transmissions are completed.

Note that the feedback of a layer index does not need to be completed right after each transmission in the forward channel. It is required only before the following key generation phase. Also note that the feedback of the index of a decodable layer is a special type of channel feedback. In particular, when considering the case when the number of fading states $L \to \infty$, the index of the highest decodable layer in time slot $m$ is equal to the fading power gain $h_{1m}$ (i.e., the public feedback $\Psi_m = h_{1m}$). For a finite level layered coding approach, the feedback of the layer index is an $L$-bit quantized version of the realization of the fading power gain. When $L = 1$, it is the ARQ feedback of ACK or NACK.

*3) Key-Generation Phase:* Once the communication phase (including feedback) is completed, both Alice and Bob can generate the secret key. Based on the feedback sequence $\mathbf{\Psi} = \mathbf{h}_1$, Alice generates a binary sequence $\mathbf{v}$ from all the messages reliably decoded by Bob based on any deterministic one-to-one mapping $g$ as

$$\mathbf{v} = g(\mathbf{W}^{[\mathcal{D}_1]}), \tag{15}$$

where $\mathbf{W}^{[\mathcal{D}_1]} = (W_1^{[\mathcal{D}_1]}, W_2^{[\mathcal{D}_1]}, \ldots, W_2^{[\mathcal{D}_1]})$ represents the set of messages successfully decoded by Bob across all layers and time slots.

Alice then looks up in the key-generation codebook for a $k$ such that $\mathbf{v} \in \mathcal{B}(k)$, and outputs $k$ as the secret key generated. Note that all those messages are decoded by Bob, and Bob can generate the same sequence $\mathbf{v}$ and the same key $k$ as Alice does. This completes the key generation.

## IV. Secrecy Key Rate

In this section, we present the secrecy key rate achieved by the broadcast approach and compare it to that achieved by using a single-level coding approach. For both approaches, we assume that the number of time slots used in the transmission over the forward channel is sufficiently large (i.e., $M \to \infty$), so that we can obtain an ergodic key rate.

[2]To be more accurate, $\mathcal{D}_1$ in $W_m^{[\mathcal{D}_1]}$ should be indexed by $m$, however, we choose to use $\mathcal{D}_1$ to simplify our notation. Throughout the paper, $W_m^{[\mathcal{D}_1]}$ is shorthand for $W_m^{[\mathcal{D}_{1m}]}$. If the subscript of $W$ is a set, then $\mathcal{D}_1$ is also indexed by the set. For example, for a set of time slots $\mathcal{M}^+ \subseteq \{1, \ldots, M\}$, we use $W_{\mathcal{M}^+}^{[\mathcal{D}_1]}$ instead of $W_{\mathcal{M}^+}^{[\mathcal{D}_{1\mathcal{M}^+}]}$ to represent all the messages decoded by Bob in $\mathcal{M}^+$. The rule is also applied to $\mathcal{D}_2$, $\mathcal{U}_1$ and $\mathcal{U}_2$. In addition, it is applied to codeword $\mathbf{X}$ and codebook $\mathcal{C}$ besides message $W$.

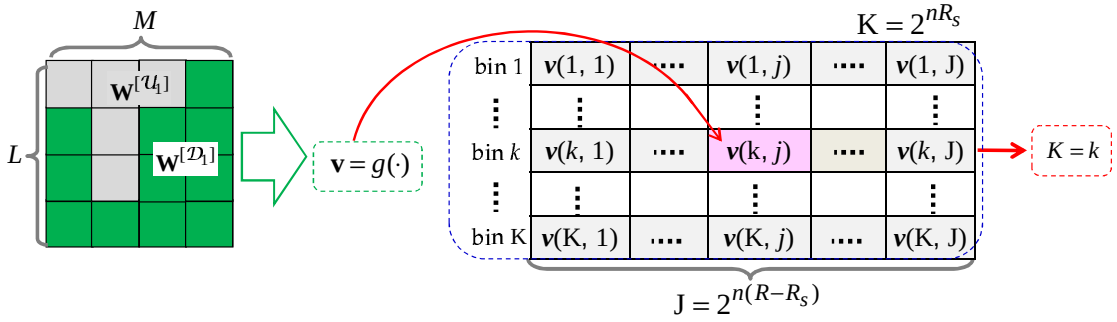


Fig. 3. Alice and Bob generate a sequence $\mathbf{v}$ from all the messages reliably decoded (across $L$ layers and $M$ time slots), look up in the key-generation codebook for a $k$ such that $\mathbf{v} \in \mathcal{B}(k)$, and output $k$ as the key.

### A. Layered-Broadcast-Coding Based Key Generation

The following result characterizes the secrecy rate when a power distribution $\rho(s)$ is given.

**Theorem 1.** For a given power distribution $\rho(s)$ over coded layers indexed by $s$, the secrecy key rate achieved by the layered-broadcast-coding based key generation scheme is

$$R_s = \int_0^{\infty} \int_0^{h_1} \Delta(h_1, h_2) dF_2(h_2) dF_1(h_1), \tag{16}$$

where $\Delta(h_1, h_2)$ is given by

$$\Delta(h_1, h_2) = \int_{h_2}^{h_1} \left[ \frac{s\rho(s)}{1+sI(s)} - \frac{h_2 \rho(s)}{1+h_2 I(s)} \right] ds \tag{17}$$

and

$$I(s) = \int_s^{\infty} \rho(u) du \quad \text{with } I(0) = P. \tag{18}$$

*Proof:* The proof can be found in Appendix A. ■

Now we discuss some insights from Theorem 1. First, $R_s$ can be written as

$$R_s = \mathbb{E}_{h_1,h_2} \left[ \tilde{\Delta}(h_1, h_2) \right], \tag{19}$$

where

$$\tilde{\Delta}(h_1, h_2) = \begin{cases} \Delta(h_1, h_2) & \text{if } h_1 > h_2 \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

The key rate $R_s$ is the average of rewards (designated by $\tilde{\Delta}(h_1, h_2)$) collected from all possible channel realizations. Positive rewards are obtained from the time slots in which Bob's channel is better than Eve's channel ($h_1 > h_2$). On the other hand, when $h_1 \leq h_2$, the reward is zero.

We can see that except for the rare case in which $h_1$ is always smaller than $h_2$, $R_s$ is positive.

Now we focus on a particular time slot $m$ in which $h_1 > h_2$, and use $\mathbf{X}_m$ to denote all layers sent in the slot. [3] As depicted in Fig. 4, $\mathbf{X}_m$ can be divided as

$$\mathbf{X}_m = \mathbf{X}_m^{[\mathcal{D}_2]} \cup \left( \mathbf{X}_m^{[\mathcal{D}_1]} \cap \mathbf{X}_m^{[\mathcal{U}_2]} \right) \cup \mathbf{X}_m^{[\mathcal{U}_1]}, \tag{21}$$

[3] $\mathbf{X}_m$ represents the set of $L$ layers in time slot $m$, and also the signal transmitted by Alice in time slot $m$, which is the superposition of all layers.

where $\mathbf{X}_m^{[\mathcal{D}_1]}$ and $\mathbf{X}_m^{[\mathcal{U}_1]}$ denote the sets of decodable and undecodable layers at Bob, respectively, and $\mathbf{X}_m^{[\mathcal{D}_2]}$ and $\mathbf{X}_m^{[\mathcal{U}_2]}$ denote the sets of decodable and undecodable layers at Eve, respectively. Note that $\mathbf{X}_m^{[\mathcal{D}_1]} \supset \mathbf{X}_m^{[\mathcal{D}_2]}$ since $h_1 > h_2$.

Both Alice and Bob can decode $\mathbf{X}_m^{[\mathcal{D}_2]}$, and neither of them can decode $\mathbf{X}_m^{[\mathcal{U}_1]}$. Therefore, a nonzero reward $\Delta(h_1, h_2)$ comes from the set of layers $\mathbf{X}_m^{[\mathcal{D}_1]} \cap \mathbf{X}_m^{[\mathcal{U}_2]}$. To show this, we rewrite (17) as

$$\Delta(h_1, h_2) = \int_{h_2}^{h_1} \frac{s\rho(s)ds}{1+sI(s)} - \int_{h_2}^{h_1} \frac{h_2 \rho(s)ds}{1+h_2 I(s)}. \tag{22}$$

The first term on the right hand side of (22) is the sum-rate decoded by Bob from $\mathbf{X}_m^{[\mathcal{D}_1]} \cap \mathbf{X}_m^{[\mathcal{U}_2]}$ (by decoding and canceling $\mathbf{X}_m^{[\mathcal{D}_2]}$ first, and treating the interference term $\mathbf{X}_m^{[\mathcal{U}_1]}$ as noise). Furthermore, the second term can be written as

$$\int_{h_2}^{h_1} \frac{h_2 \rho(s)ds}{1+h_2 I(s)} = \log \left( 1 + \frac{h_2 \left[ I(h_2) - I(h_1) \right]}{1+h_2 I(h_1)} \right). \tag{23}$$

By noticing that $I(h_2) - I(h_1)$ is the total power used for the layers $\mathbf{X}_m^{[\mathcal{D}_1]} \cap \mathbf{X}_m^{[\mathcal{U}_2]}$, and $I(h_1)$ is the total power used for the layers $\mathbf{X}_m^{[\mathcal{U}_1]}$, (23) gives the rate of information that Eve can possibly deduce from $\mathbf{X}_m^{[\mathcal{D}_1]} \cap \mathbf{X}_m^{[\mathcal{U}_2]}$ through her channel with power gain $h_2$.

An interesting finding here is that what the best Eve can do is to treat the interference term $\mathbf{X}_m^{[\mathcal{U}_1]}$ as noise (as Bob does) with the total noise power $1+h_2 I(h_1)$, and therefore cannot benefit from the structure of interference either. Due to the absence of CSI at the transmitter during the transmission in the forward channel , the layered broadcast coding strategy creates a medium with interference, in which the undecodable layers play the role of *self-interference*. We remark that this is a special case of secret communication over a medium with interference as discussed in [27].

### B. Single-Level-Coding Based Key Generation

When single-level coding is used, self-interference does not occur. Alice uses a codebook with a single coding rate in the forward transmission. Bob uses ARQ feedback to tell

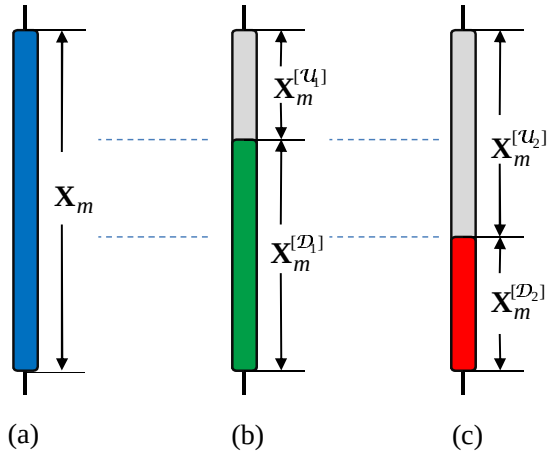


Fig. 4. (a) Coded layers sent by Alice, (b) decodable and undecodable layers for Bob, and (c) decodable and undecodable layers for Eve, in time slot $m$ with the channel gains $h_1 > h_2$.

Alice whether the decoding is successful or has failed. In this case, the following secrecy key rate can be achieved.

**Lemma 1.** [30, Theorem 1] The secrecy key rate of a single-level-coding based scheme is given by

$$R_s^{[1]} = \Pr\left[R^{[1]} \le \log(1+h_1P)\right] \times \mathbb{E}_{h_2}\left[R^{[1]} - \log\left(1+h_2P\right)\right]^+, \quad (24)$$

where $R^{[1]}$ is the coding rate of the single-level codebook.

This key rate $R_s^{[1]}$ still has the interpretation of the average of rewards (designated by $\tilde{\Delta}_1(h_1,h_2)$) collected from all possible channel realizations. That is, $R_s^{[1]}$ can be written as

$$R_s^{[1]} = \mathbb{E}_{h_1,h_2}\left[\tilde{\Delta}_1(h_1,h_2)\right], \quad (25)$$

where

$$\tilde{\Delta}_1(h_1,h_2) = \begin{cases} R^{[1]} - \log(1+h_2P) & \text{if } h_1 \ge \frac{\exp(R^{[1]})-1}{P} > h_2 \\ 0 & \text{otherwise.} \end{cases} \quad (26)$$

### C. Comparisons and Discussions

The advantage of the layered-broadcast-coding (LBC) based approach over the single-level-coding based approach (SLC) can be readily observed by comparing the reward functions given by (20) and (26). First, in LBC, a positive reward is obtained from the set of channel pairs $\mathcal{P} = \{(h_1,h_2) : h_1 > h_2\}$; while in SLC, it is obtained from the channel set $\mathcal{P}' = \{(h_1,h_2) : h_1 \ge \frac{1}{P}(e^{R^{[1]}}-1) > h_2\}$. It is obvious that $\mathcal{P} \supset \mathcal{P}'$, which means there are more time slots that contribute to the secrecy key generation for LBC than for SLC. Second, the coding rate $R^{[1]}$ for SLC has to be carefully chosen in order to balance between obtaining a larger value of reward in a time slot (by increasing $R^{[1]}$) and making more time slots contribute to the key generation (by decreasing $R^{[1]}$); while in LBC, the reward is gained in each time slot adaptively based on the random channel realizations. Finally and importantly, in SLC, Eve can deduce the information at the rate of $\log(1+h_2P)$ with a channel gain $h_2$. This is the loss of rate in order to keep the key secret from Eve. In LBC, however, Eve deduces less information as given by (23) due to the interference power (the total power of undecodable layers). The self-interference plays an important role for decreasing Eve's capability of eavesdropping.

Hence, although the single-level-coding based approach has lower decoding complexity, and requires less feedback (only 1-bit per time slot), it is sub-optimal in general (when feedback of multiple bits is allowed). By all means, the single-level coding scheme can be considered as a special case of a layered-broadcast-coding based scheme, in which all power is allocated to a single layer. It serves as a baseline scheme and further motivates us to find the best power distribution for optimizing the layered-broadcast-coding scheme.

## V. Optimal Power Distribution

In this section, we derive the optimal distribution of power over coded layers for our broadcast approach. The secrecy rate given by (16) is hard to evaluate and optimize due to the three-dimensional integrals. After some steps of derivations, we have an alternative form given as follows:

**Lemma 2.** The secrecy key rate given by (16) is equivalent to

$$R_s = \max_{I(x)} \int_0^\infty \left[1-F_1(x)\right]\rho(x)\left[\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2}\right]dx, \quad (27)$$

with the constraint $I(0)=P$, and $\rho(x) = -dI(x)/dx$.

*Proof:* The proof can be found in Appendix E. ■

### A. Optimal Interference Distribution

In certain cases, optimization of $R_s$ with respect to the power distribution $\rho(x)$, or equivalently, the interference distribution $I(x)$, under the power constraint $P$ can be found by using the calculus of variations. First, we define the functional of (27) as

$$L\left(x, I(x), I'(x)\right) = -\left[1-F_1(x)\right]I'(x)\left[\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2}\right].$$

A necessary condition for a maximum of the integral of $L(x,I(x),I'(x))$ over $x$ is a zero variation of the functional. By solving the associated Eüler-Lagrangian equation [32] given as

$$\frac{\partial L}{\partial I} - \frac{d}{dx}\left(\frac{\partial L}{\partial I'}\right) = 0, \quad (28)$$

we have the following characterization for the optimal $I(x)$.

**Theorem 2.** A necessary condition for optimizing $I(x)$ in order to maximize the secrecy rate given by (27) is to choose $I(x)$ to satisfy

$$\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2} = \frac{[1-F_1(x)]\,F_2(x)}{f_1(x)\,[1+xI(x)]^2}, \tag{29}$$

where $I(x)=0$ when $x<x_0$ or $x\geq x_1$. Here, $x_0$ and $x_1$ can be found by setting $I(x_0)=P$ and $I(x_1)=0$ in (29).

*Proof:* The proof can be found in Appendix F. ■

In general, numerical computation is needed for solving (29) in order to obtain the optimal interference distribution $I(x)$. For some special CDFs $F_2(x)$, an analytical form of $I(x)$ is possible if the integral in (29) can be evaluated in a closed form.

In the following, we consider two of such special cases:

*1) Non-Fading Alice-Eve Channel:* If the Alice-Eve channel is constant with channel power gain $x^*$, the CDF $F_2(x)$ is $F_2(x)=\mu(x-x^*)$, where $\mu(x)$ represents a unit step function. In this case, the optimal interference distribution is given by

$$I(x) = \frac{1-F_1(x)-(x-x^*)f_1(x)}{x(x-x^*)f_1(x)-x^*\,[1-F_1(x)]}, \tag{30}$$

which can be easily shown from (29).

*2) Non-Secret Layered Transmission:* If key-generation is not considered and it is desired to find the optimal $I(x)$ to maximize the average reliably decodable rate at Bob in the non-secret layered transmission, this can be done by assuming $x^*=0$ in (30). In this case, we have

$$I(x) = \frac{1-F_1(x)}{x^2 f_1(x)} - \frac{1}{x}, \tag{31}$$

which is consistent with the result given in [29].

### B. Secrecy Key Rate With Optimal Power Distribution

Finally, we have the following secrecy key rate under the optimal power distribution.

**Corollary 1.** When the optimal power distribution is used, the following secrecy key rate is achieved:

$$R_s = \int_{x_0}^{x_1} \frac{-[1-F_1(x)]^2\,F_2(x)dI(x)}{f_1(x)[1+xI(x)]^2}, \tag{32}$$

where $I(x)$ and $(x_0,x_1)$ are found from the condition given by Theorem 2.

*Proof:* The proof is straightforward by combining Lemma 2 and Theorem 2. ■

## VI. A Rayleigh Fading Channel

In this section, we assume Rayleigh fading for both Alive-Bob and Alice-Eve channels. The fading gains $h_t$ are exponentially distributed with means $\lambda_t$ for $t=1,2$. That is, the PDFs of the fading gain $h_t$ are

$$f_t(s) = \begin{cases} \frac{1}{\lambda_t}\exp\left(-\frac{s}{\lambda_t}\right) & \text{if } s\geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{33}$$

for $t=1,2$ and the CDFs are

$$F_t(s) = \begin{cases} 1-\exp\left(-\frac{s}{\lambda_t}\right) & \text{if } s\geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

### A. Single-Level-Coding Approach

For comparison, we first calculate the secrecy key rate when single-level coding is used. As shown in Appendix G, the secrecy rate is

$$R_s^{[1]} = \max_{R^{[1]}\geq 0} \quad \exp\left(-\frac{e^{R^{[1]}}-1}{\lambda_1 P}\right)\times \left\{R^{[1]} - \exp\left(\frac{1}{\lambda_2 P}\right)\left[E_i\left(\frac{e^{R^{[1]}}}{\lambda_2 P}\right) - E_i\left(\frac{1}{\lambda_2 P}\right)\right]\right\}, \tag{35}$$

where $E_i(x)=\int_x^\infty[\exp(-t)/t]dt$ is the exponential integral function. It can be verified that the above function is concave with respect to $R^{[1]}$ and thus has a unique maximum, which can be searched numerically.

### B. Layered-Coding Approach

According to (32), the secrecy rate with layered coding under the optimal power control is computed numerically by evaluating

$$R_s = \lambda_1 \int_{x_0}^{x_1} \frac{\exp(-x/\lambda_1)\,[\exp(-x/\lambda_2)-1]}{[1+xI(x)]^2} dI(x),$$

where the optimal interference distribution $I(x)$ and boundary points $x_0$ and $x_1$ can be found according to Lemma 2 as follows.

*1) Interference Distribution $I(x)$:* As shown in Appendix H, we have

$$\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2} = \frac{\exp\left(-x/\lambda_2\right)-1}{I(x)\,[1+xI(x)]} + \frac{\exp\left(1/\lambda_2 I(x)\right)}{\lambda_2 I^2(x)}\left[E_i\left(\frac{1}{\lambda_2 I(x)}\right) - E_i\left(\frac{1+xI(x)}{\lambda_2 I(x)}\right)\right]. \tag{36}$$

We also have

$$\frac{[1-F_1(x)]\,F_2(x)}{f_1(x)\,[1+xI(x)]^2} = \frac{\lambda_1\,[1-\exp(-x/\lambda_2)]}{[1+xI(x)]^2}. \tag{37}$$

Therefore, we can show after some steps of arrangements that $I(x)$ is found by solving

$$E_i\left(\frac{1}{\lambda_2 I(x)}\right) - E_i\left(\frac{1+xI(x)}{\lambda_2 I(x)}\right) = \frac{\lambda_2 I(x)[1+\lambda_1 I(x)]}{[1+xI(x)]^2} \times \left[\exp\left(-\frac{1}{\lambda_2 I(x)}\right) - \exp\left(-\frac{1+xI(x)}{\lambda_2 I(x)}\right)\right]. \tag{38}$$

*2) Boundary Points $x_0$ and $x_1$:* We needs to find the boundary points $x_0$ and $x_1$ to meet the constraints that

$$I(x_0) = P \quad \text{and} \quad I(x_1) = 0.$$

By letting $I(x_0) = P$ in (38), we can solve the equation for $x_0$. However, $x_1$ cannot be solved by this means since we cannot let $I(x_1) = 0$ in (38). Instead, we let $I(x_1) = 0$ in (29) and find that

$$\int_0^{x_1} F_2(y)dy = x_1 + \lambda_2 \left[\exp(-x_1/\lambda_2) - 1\right],$$

and

$$\frac{[1 - F_1(x_1)]\, F_2(x_1)}{f_1(x_1)} = \lambda_1 \left[1 - \exp\left(-x_1/\lambda_2\right)\right].$$

Therefore, $x_1$ can be found by solving the following equation:

$$x_1 + (\lambda_1 + \lambda_2)\left[\exp\left(-\frac{x_1}{\lambda_2}\right) - 1\right] = 0.$$

Interestingly, $x_1$ depends only on the channel statistics (characterized by $\lambda_1$ and $\lambda_2$ for the Rayleigh fading channels) and not on the power constraint $P$. Note that no power will be allocated to a layer with its index higher than $x_1$ (however, it is possible that some layers lower than $x_1$ still have zero power allocation, as shown in the numerical example). Finally, we remark that every equation discussed in this section has a unique solution after excluding a trivial solution $0$.

### C. Numerical Examples

Now we show some numerical examples on the achievable secrecy-key rates and the optimal power distribution $\rho(s)$. We consider the symmetric Rayleigh fading channel defined by (33) with $\lambda_1 = \lambda_2 = 1$.

Fig. 5 compares the secrecy key rates achieved by the layered-coding and single-level-coding based schemes (both optimized). We also compare them with the secrecy rate when perfect and noncausal CSI of the Alice-Bob channel is available to Alice. In this case, Alice is able to adapt its transmission rate based on the CSI at each time slot. We still assume a short-term power constraint and thus Alice does not adapt power in contrast to the scheme given by [12]. Without CSI at Alice, the secrecy key rate achieved by the layered-coding based scheme is significantly higher. This shows the benefit of the broadcast approach due to the introduction of self-interference in transmission.

Fig. 6 shows the optimal power distribution over coded layers. A trend is that more power is distributed to lower layers as the total transmit power $P$ becomes larger. In general, the optimal power distribution does not concentrate much on a certain layer (or a small set of layers), especially when $P$ is large. We also compare the optimal power distribution for maximizing the secrecy key rate in key-generation and that for maximizing the average reliably decodable rate at Bob in non-secret transmission. With different power constraints, the power distributions for non-secret transmission are on the same curve but have different boundary points, which is different from the case for key generation. Also, when the total transmit power exceeds a certain threshold, the power distribution for key generation is more concentrated over higher layers (as shown for the cases of $P = 5$ and $P = 20$); while the opposite can be observed when $P$ is small (as shown for the case of $P = 1$ in Fig. 6.)

## VII. Conclusions

In this paper, we have introduced a broadcast approach for secret-key generation over slow-fading channels based on layered broadcast coding. We have considered a model in which Alice attempts to share a key with Bob while keeping the key secret from Eve. Both Alice-Bob and Alice-Eve channels are assumed to undergo slow fading, and perfect CSI is assumed to be known only at the receivers during the transmission. Layered coding facilitates adapting the reliably decoded rate at Bob to the actual channel state without CSI available at Alice. The index of a reliably decoded layer is sent back to Alice via an authenticated, public and error-free channel, which is exploited by Alice and Bob to generate the secret key. We have derived the achievable secrecy key rate and characterized the optimal power distribution over coded layers. Our theoretical and numerical results have shown that the broadcast approach outperforms the single-level-coding based approach significantly, which establishes the important role of introducing self-interference in facilitating secret-key generation over slow-fading channels when transmit CSI is not available.

## Appendix A
## Proof of Theorem 1

Let us first consider the $L$-state fading wiretap channel defined by Definition 2. We have the following result.

**Lemma A.1.** For the $L$-state fading wiretap channel defined by Definition 2, the following key-rate is achievable:

$$R_s = \sum_{l_1} \sum_{l_2 < l_1} \Pr\left(h_1 = h^{[l_1]}, h_2 = h^{[l_2]}\right) \times \sum_{l=l_2+1}^{l_1} \left[r^{[l]} - \log\left(1 + \frac{h^{[l_2]} p^{[l]}}{1 + h^{[l_2]} \sum_{i=l+1}^{L} p^{[i]}}\right)\right], \tag{39}$$

where we assume that $\{h^{[1]} \le h^{[2]} \le \cdots \le h^{[L]}\}$ and $r^{[l]}$ is given by

$$r^{[l]} = \log\left(1 + \frac{h^{[l]} p^{[l]}}{1 + h^{[l]} \sum_{i=l+1}^{L} p^{[i]}}\right). \tag{40}$$

*Proof:* We relegate the proof of Lemma A.1 to Appendix B. ∎

It is easy to observe that the result given by Theorem 1 is a continuous version of Lemma A.1 (as $L \to \infty$), and can be shown by following some standard steps in a

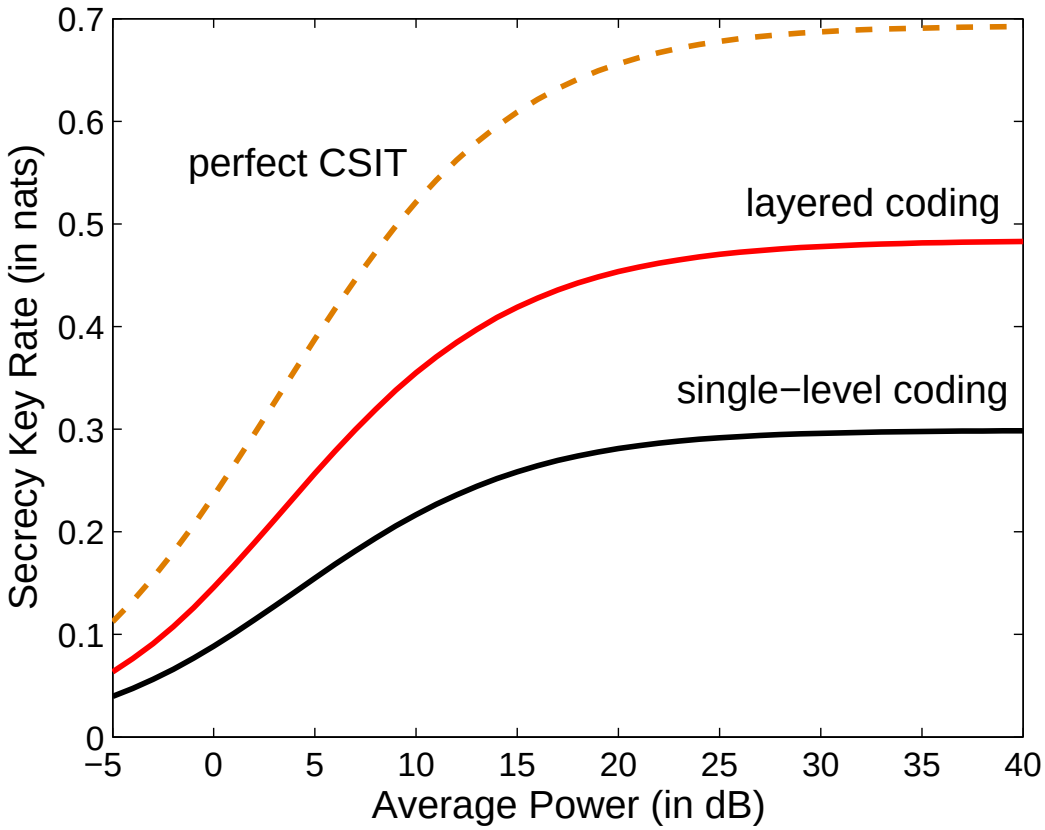


Fig. 5. Secrecy key rates achievable for the layered-coding-based approach, the single-level-coding-based approach, and when perfect CSIT is available at Alice noncausually.

straightforward manner. We omit these steps and next prove Lemma A.1 only.

## Appendix B
## Proof of Lemma A.1

### A. Secret-Key Generation For The $L$-State Fading Wiretap Channel

The key-generation scheme for the $L$-state fading wiretap channel is similar to the scheme outlined in Section III-C. The encoding and decoding in the communication phase have been discussed in Section III-B. To proceed with the key generation phase, we will use the following notation (some of which has been explained previously but is repeated here for ease of reference).

Let $W_m = W_m^{[1:L]}$ represent the set of messages sent by Alice at the $m$-th time slot and $W_m^{[l]}$ represents the message sent at the $l$-th layer. At Bob, the reliably decoded message set at the $m$-th time slot is denoted by $W_m^{[\mathcal{D}_1]}$ and the undecodable message set is denoted by $W_m^{[\mathcal{U}_1]}$. At Eve, similarly, the reliably decoded message set is denoted by $W_m^{[\mathcal{D}_2]}$ and the undecodable message set is $W_m^{[\mathcal{U}_2]}$. We use $\mathbf{W} = (W_1, W_2, \ldots, W_M)$ to represent the set of messages sent over all $M$ time slots. Similarly, $\mathbf{W}^{[\mathcal{D}_t]} = (W_1^{[\mathcal{D}_t]}, W_2^{[\mathcal{D}_t]}, \ldots, W_M^{[\mathcal{D}_t]})$ and $\mathbf{W}^{[\mathcal{U}_t]} = (W_1^{[\mathcal{U}_t]}, W_2^{[\mathcal{U}_t]}, \ldots, W_M^{[\mathcal{U}_t]})$ are defined for $t = 1, 2$.

We use $\mathbf{X}_m = \mathbf{X}_m^{[l:L]}$ to represent the set of codewords sent in the $m$-th time slot, $\mathbf{X}_m^{[\mathcal{D}_t]}$ and $\mathbf{X}_m^{[\mathcal{U}_t]}$ (for $t = 1, 2$) to represent the sets of reliably decoded, and undecodable layers, respectively. Furthermore, $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_M)$, $\mathbf{X}^{[\mathcal{D}_t]} = (\mathbf{X}_1^{[\mathcal{D}_t]}, \mathbf{X}_2^{[\mathcal{D}_t]}, \ldots, \mathbf{X}_M^{[\mathcal{D}_t]})$ and $\mathbf{X}^{[\mathcal{U}_t]} = (\mathbf{X}_1^{[\mathcal{U}_t]}, \mathbf{X}_2^{[\mathcal{U}_t]}, \ldots, \mathbf{X}_M^{[\mathcal{U}_t]})$ are the set, reliably decoded set, and undecodable set of codewords, respectively, over all $M$ time slots. In addition, $\mathbf{Y}_1 = (\mathbf{Y}_{1,1}, \mathbf{Y}_{1,2}, \ldots, \mathbf{Y}_{1,M})$ and $\mathbf{Y}_2 = (\mathbf{Y}_{2,1}, \mathbf{Y}_{2,2}, \ldots, \mathbf{Y}_{2,M})$ are the signals observed by Bob and Eve, respectively, over all $M$ time slots.

In the key generation phase, two parameters of the key generation codebook are $R$ and $R_s$. For the $L$-state fading wiretap channel, $R_s$ is given by (39) and $R$ is given by

$$R = \sum_{l=1}^{L} \Pr\left(h_1 = h^{[l]}\right) \left(\sum_{i=1}^{l} r^{[i]}\right), \tag{41}$$

where $r^{[i]}$ is given by (40).

### B. Genie-Leaked Information

In the communication phase, we assume that the message conveyed by each layer is chosen independently of those at all other layers and uniformly at random. That is, at

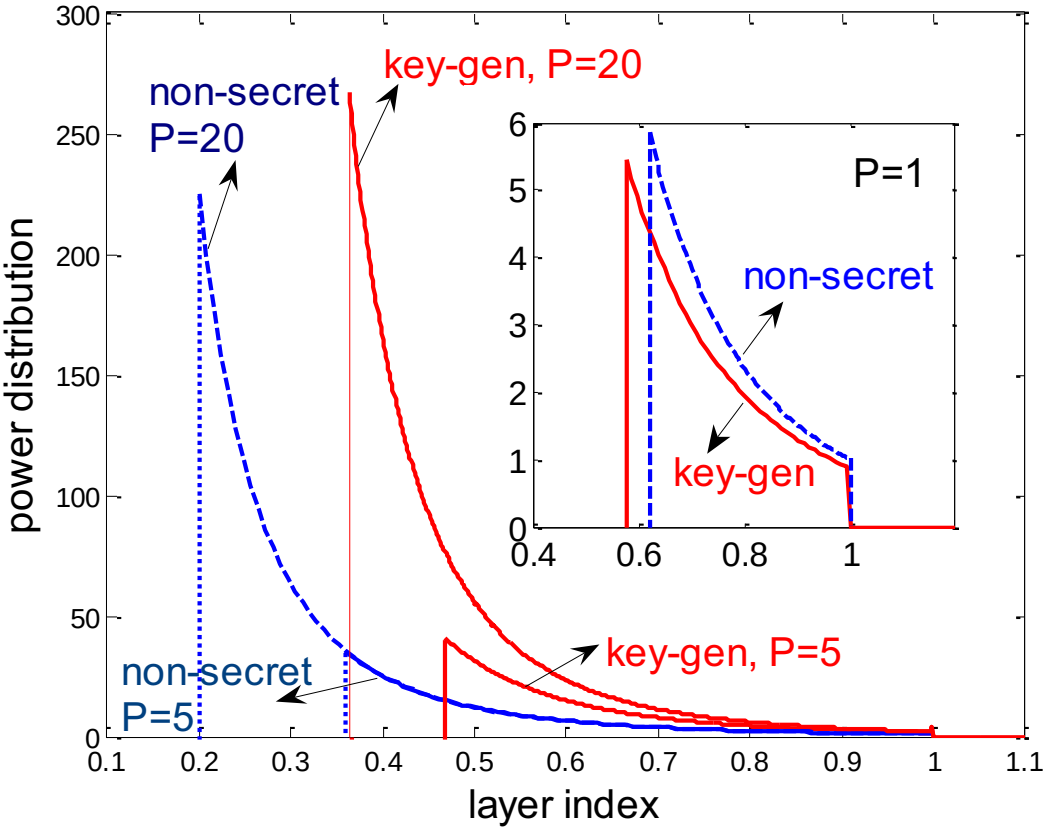


Fig. 6. Optimal power distributions for maximizing the secrecy key rate in key-generation ("key-gen") and for maximizing the average reliably decodable rate at Bob in non-secret transmission ("non-secret") when the normalized transmit power is $P = 1, 5, 20$.

time slot $m$, the message $W_m^{[l]}$ sent by the $l$-th layer, is randomly and uniformly selected from $\{1, 2, \ldots, 2^{Nr^{[l]}}\}$. One can always assume that the random message is generated through a two-step procedure: first, two messages $\hat{W}_m^{[l]}$ and $\check{W}_m^{[l]}$ are selected randomly and independently, where $\hat{W}_m^{[l]} \in \{1, \ldots, 2^{N\hat{r}_m^{[l]}}\}$ and $\check{W}_m^{[l]} \in \{1, \ldots, 2^{N\check{r}_m^{[l]}})\}$, where $\check{r}_m^{[l]} = r^{[l]} - \hat{r}_m^{[l]}$; Then, message $W_m^{[l]} = \hat{W}_m^{[l]} \times \check{W}_m^{[l]}$ is formed.

Note that this procedure is assumed only for facilitating the proof and is not actually required for encoding. In fact, $\hat{r}_m^{[l]}$ can be any value as long as $0 \le \hat{r}_m^{[l]} \le r^{[l]}$. For example, we can assume the following value for $\hat{r}_m^{[l]}$:

- if $1 \le l \le l_{1m}$ (i.e., $l \in \mathcal{D}_{1m}$),

$$\hat{r}_m^{[l]} = r^{[l]}; \tag{42}$$

- otherwise,

$$\hat{r}_m^{[l]} = \min\left\{r^{[l]}, \log\left(1 + \frac{h_{2m}p^{[l]}}{1 + h_{2m}\sum_{i=l+1}^{L} p^{[i]}}\right)\right\}, \tag{43}$$

where $l_{1m}$ is the feedback layer index (i.e., the highest index of the decodable layers at Bob) in time slot $m$. Again, the feedback and channel information are not needed during the transmission since the two-step procedure is not actually executed.

Following the partitioning of messages, we have $\hat{W}_m^{[\mathcal{D}_1]} = W_m^{[\mathcal{D}_1]}$, $\check{W}_m^{[\mathcal{D}_1]} = \emptyset$, and $W_m^{[\mathcal{U}_1]} = \hat{W}_m^{[\mathcal{U}_1]} \times \check{W}_m^{[\mathcal{U}_1]}$. Hence, $W_m$ is decomposed as $W_m = W_m^{[\mathcal{D}_1]} \times \hat{W}_m^{[\mathcal{U}_1]} \times \check{W}_m^{[\mathcal{U}_1]}$. By letting $\hat{\mathbf{W}}^{[\mathcal{U}_1]} = (\hat{W}_1^{[\mathcal{U}_1]}, \hat{W}_2^{[\mathcal{U}_1]}, \ldots, \hat{W}_M^{[\mathcal{U}_1]})$ and $\check{\mathbf{W}}^{[\mathcal{U}_1]} = (\check{W}_1^{[\mathcal{U}_1]}, \check{W}_2^{[\mathcal{U}_1]}, \ldots, \check{W}_M^{[\mathcal{U}_1]})$, we have $\mathbf{W} = \mathbf{W}^{[\mathcal{D}_1]} \times \hat{\mathbf{W}}^{[\mathcal{U}_1]} \times \check{\mathbf{W}}^{[\mathcal{U}_1]}$ correspondingly.

We assume that there is a genie who gives the message set $\check{\mathbf{W}}^{[\mathcal{U}_1]}$ to Eve. This is a useful step to enable us to give a bound on the equivocation rate with respect to the key $K$ at Eve.

One might wonder if this genie-leaked information benefits Eve and eventually reduces the achievable key rate. In Fig. 7, we illustrate that the genie-leaked information does not benefit Eve. Here, let us consider a special $L$-state fading wiretap channel for which $L = 3$ and the support of both Alice-Bob and Alice-Eve channel gains is $\{h^{[1]}, h^{[2]}, h^{[3]}\}$. It is easy to see that $\check{W}_m^{[\mathcal{U}_1]} \neq \emptyset$ if and only if $h_{1m} = h^{[2]}$ and $h_{2m} = h^{[1]}$ for a time slot $m$. Therefore, we can focus on such a time slot. We have $\mathcal{D}_{1m} = \{1, 2\}$, $\mathcal{U}_{1m} = \{3\}$, $\mathcal{D}_{2m} = \{1\}$, and $\mathcal{U}_{2m} = \{2, 3\}$.

$\mathbf{X}_m^{[1]}$ is decoded and subtracted by both Alice and Bob from their received signals. Therefore, we consider only $\mathbf{X}_m^{[2]}$ and $\mathbf{X}_m^{[3]}$, where $\mathbf{X}_m^{[2]}$ contributes to the key generation and Eve tries to deduce information on $\mathbf{X}_m^{[2]}$, while $\mathbf{X}_m^{[3]}$

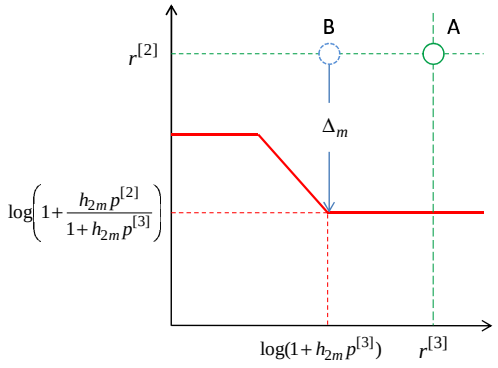


Fig. 7. An illustrative example to show that the genie-leaked information does not benefit Eve.

plays the role of interference. Fig. 7 shows the rate of information that Eve can deduce on $\mathbf{X}_m^{[2]}$ versus the rate of interference codebook. (The rate region resembles that of a multiple access channel. Some related discussion can be found in [27].)

Eve uses the genie-leaked information to reduce the rate of interference codebook. To achieve this, Eve uses $\breve{W}_m^{[3]}$ to obtain a thinned codebook $\mathcal{C}^{[3]}(\breve{W}_m^{[3]})$. That is, among all the codewords in the original codebook $\mathcal{C}^{[3]}$, i.e. only the ones corresponding to $\breve{W}_m^{[3]}$ are kept and the rest are eliminated. However, if the side information is given properly, Eve does not benefit from the genie. As shown in Fig. 7, the side information does not help Eve's eavesdropping if

$$\breve{r}_m^{[3]} \le r^{[3]} - \log\left(1+h_{2m}p^{[3]}\right).$$

Under this condition, the pair of coding rates of $\mathcal{C}^{[2]}$ and $\mathcal{C}^{[3]}(\breve{W}_m^{[3]})$ is represented by any point on the line segment from A to B. A reward of

$$\Delta_m = r^{[2]} - \log\left(1+\frac{h_{2m}p^{[2]}}{1+h_{2m}p^{[3]}}\right)$$

is collected from time slot $m$ in contributing to the key generation.

### C. Equivocation Calculation

Now, we are ready to compute the equivocation rate with respect to the key $K$ at Eve:

$$\begin{aligned}
&H(K|\mathbf{Y}_2,\mathbf{\Psi},\mathbf{h}_2)\\
&\ge H(K|\mathbf{Y}_2,\mathbf{\Psi},\mathbf{h}_1,\mathbf{h}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]}) &(44)\\
&= H(K|\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2) &(45)\\
&= H(K,\mathbf{Y}_2,\mathbf{X}|\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2) - H(\mathbf{Y}_2|\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2)\\
&\quad - H(\mathbf{X}|\mathbf{Y}_2,K,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2)\\
&\ge H(\mathbf{X}|\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2)\\
&\quad - H(\mathbf{X}|\mathbf{Y}_2,K,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2), &(46)
\end{aligned}$$

where (44) is from the property that conditioning reduces entropy, (45) is due to the fact that $\mathbf{\Psi}$ is a deterministic function of $\mathbf{h}_1$ and $\mathbf{Y}_2$.

As shown in Appendix C and D, the two terms in (46) can be bounded as in the following,

$$H(\mathbf{X}|\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2) \ge n(R_s-\delta_{N,M}), \quad (47)$$

and

$$H(\mathbf{X}|\mathbf{Y}_2,K,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2) \le n\delta'_{N,M}, \quad (48)$$

where $\delta_{N,M},\delta'_{N,M}\to 0$ when $N,M\to\infty$.

By combining (46), (47) and (48), we have

$$nR_e = H(K|\mathbf{Y}_2,\mathbf{\Psi},\mathbf{h}_2) \ge n(R_s-\delta), \quad (49)$$

which gives the perfect secrecy requirement that is

$$R_e \ge R_s - \delta,$$

where $\delta\to 0$ as $n\to\infty$ (actually $N,M\to\infty$). Hence, we complete the proof.

## Appendix C
## Proof of (47)

First, let us denote

$$E_1 \triangleq H(\mathbf{X}|\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2).$$

Due to independent coding at each time slot during forward transmission, we have

$$\begin{aligned}
E_1 &= H(\mathbf{X},\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2) - H(\mathbf{Y}_2,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2)\\
&= \sum_{m=1}^{M} H(\mathbf{X}_m,\mathbf{Y}_{2m},\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m})\\
&\quad - \sum_{m=1}^{M} H(\mathbf{Y}_{2m},\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m})\\
&= \sum_{m=1}^{M} H(\mathbf{X}_m|\mathbf{Y}_{2m},\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m}).
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
E_1 &\ge \sum_{m\in\mathcal{M}^+} H(\mathbf{X}_m|\mathbf{Y}_{2m},\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m}) &(50)\\
&= \sum_{m\in\mathcal{M}^+} H(\mathbf{X}_m|\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m})\\
&\quad + H(\mathbf{Y}_{2m}|\mathbf{X}_m,\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m})\\
&\quad - H(\mathbf{Y}_{2m}|\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m})\\
&= \sum_{m\in\mathcal{M}^+} H(\mathbf{X}_m|\breve{W}_m^{[\mathcal{U}_1]}) + H(\mathbf{Y}_{2m}|\mathbf{X}_m,h_{2m})\\
&\quad - H(\mathbf{Y}_{2m}|\breve{W}_m^{[\mathcal{U}_1]},h_{1m},h_{2m}) &(51)\\
&\ge \sum_{m\in\mathcal{M}^+} H(\mathbf{X}_m|\breve{W}_m^{[\mathcal{U}_1]}) + H(\mathbf{Y}_{2m}|\mathbf{X}_m,h_{2m})\\
&\quad - H(\mathbf{Y}_{2m}|h_{2m}) &(52)\\
&\ge \sum_{m\in\mathcal{M}^+} H(\mathbf{X}_m|\breve{W}_m^{[\mathcal{U}_1]}) - I(\mathbf{X}_m;\mathbf{Y}_{2m}|h_{2m}) &(53)
\end{aligned}$$

where $\mathcal{M}^+ = \{m|m\in\{1,\dots,M\},h_{1m}\ge h_{2m}\}$ is the set of time slots in which Alice-Bob channel is better than

Alice-Eve channel), (50) follows from the property that entropy is non-negative, (51) follows from the property that $\check{W}_m^{[\mathcal{U}_1]} \leftrightarrow \mathbf{X}_m \leftrightarrow \mathbf{Y}_{2m}$ forms a Markov chain, and (52) follows from the property that conditioning reduces entropy.

To bound (53) further, we have

$$\begin{aligned} &H(\mathbf{X}_m|\check{W}_m^{[\mathcal{U}_1]}) \\ &= N\left(\sum_{l=1}^{L}\hat{r}_m^{[l]}\right) \\ &= N\left[\sum_{l=1}^{l_{1m}} r^{[l]} + \sum_{l=l_{1m}+1}^{L}\log\left(1+\frac{h_{2m}p^{[l]}}{1+h_{2m}\sum_{i=l+1}^{L}p^{[i]}}\right)\right] \quad (54) \\ &= N\left[\sum_{l=1}^{l_{1m}} r^{[l]} + \log\left(1+h_{2m}\sum_{l=l_{1m}+1}^{L}p^{[l]}\right)\right] \quad (55) \end{aligned}$$

where $l_{1m}$ denotes the index of the highest decodable layer at Bob in time slot $m$, and (54) follows from (43). We also have

$$\begin{aligned} &I\left(\mathbf{X}_m;\mathbf{Y}_{2m}|h_{2m}\right) \\ &= I\left(\mathbf{X}_m^{[\mathcal{D}_2]},\mathbf{X}_m^{[\mathcal{U}_2]};\mathbf{Y}_{2m}|h_{2m}\right) \\ &= I\left(\mathbf{X}_m^{[\mathcal{D}_2]};\mathbf{Y}_{2m}|h_{2m}\right) + I\left(\mathbf{X}_m^{[\mathcal{U}_2]};\mathbf{Y}_{2m}|\mathbf{X}_m^{[\mathcal{D}_2]},h_{2m}\right) \\ &\le H(\mathbf{X}_m^{[\mathcal{D}_2]}) + I\left(\mathbf{X}_m^{[\mathcal{U}_2]};\mathbf{Y}_{2m}|\mathbf{X}_m^{[\mathcal{D}_2]},h_{2m}\right) \\ &\le N\left[\sum_{l=1}^{l_{2m}} r^{[l]} + \log\left(1+h_{2m}\sum_{l=l_{2m}+1}^{L}p^{[l]}\right)+\delta_1\right], \quad (56) \end{aligned}$$

where $l_{2m}$ denotes the index of the highest decodable layer at Eve in time slot $m$, and $\delta_1 \to 0$ as $N \to \infty$.

Combining (53), (55), and (56), we have

$$\begin{aligned} E_1 &\ge N\left\{\sum_{m\in\mathcal{M}^+}\left[\sum_{l=l_{2m}+1}^{l_{1m}} r^{[l]} - \right.\right. \\ &\left.\left. \log\left(1+\frac{h_{2m}\sum_{l=l_{2m}+1}^{l_{1m}}p^{[l]}}{1+h_{2m}\sum_{l=l_{1m}+1}^{L}p^{[l]}}\right)-\delta_1\right]\right\} \\ &= N\sum_{l_1}\sum_{l_2<l_1}\#\left(h_1=h^{[l_1]},h_2=h^{[l_2]}\right)\times \\ &\left[\sum_{l=l_2+1}^{l_1} r^{[l]} - \log\left(1+\frac{h^{[l_2]}\sum_{l=l_2+1}^{l_1}p^{[l]}}{1+h^{[l_2]}\sum_{l=l_1+1}^{L}p^{[l]}}\right)-\delta_1\right] \\ &= N\sum_{l_1}\sum_{l_2<l_1}\#\left(h_1=h^{[l_1]},h_2=h^{[l_2]}\right)\times \\ &\left\{\sum_{l=l_2+1}^{l_1}\left[r^{[l]} - \log\left(1+\frac{h^{[l_2]}p^{[l]}}{1+h^{[l_2]}\sum_{i=l+1}^{L}p^{[i]}}\right)\right]-\delta_1\right\}, \end{aligned}$$

where $\#\left(h_1=h^{[l_1]},h_2=h^{[l_2]}\right)$ denotes the number of time slots (out of $M$ slots) that $h_1=h^{[l_1]}$ and $h_2=h^{[l_2]}$.

When $M\to\infty$, we have

$$\begin{aligned} E_1 &\ge N\sum_{l_1}\sum_{l_2<l_1} M\left[\Pr\left(h_1=h^{[l_1]},h_2=h^{[l_2]}\right)-\delta_1'\right] \\ &\times\left\{\sum_{l=l_2+1}^{l_1}\left[r^{[l]}-\log\left(1+\frac{h^{[l_2]}p^{[l]}}{1+h^{[l_2]}\sum_{i=l+1}^{L}p^{[i]}}\right)\right]-\delta_1\right\} \\ &= n(R_s-\delta_2), \quad (57) \end{aligned}$$

where $\delta_2\to 0$ when $N\to\infty$ and $M\to\infty$.

## Appendix D
## Proof of (48)

First, we denote

$$E_2 \triangleq H(\mathbf{X}|\mathbf{Y}_2,K,\breve{\mathbf{W}}^{[\mathcal{U}_1]},\mathbf{h}_1,\mathbf{h}_2).$$

To give a bound on $E_2$, we consider Eve's decoding of $\mathbf{X}$, i.e., the codewords sent over all $L$ layers and $M$ time slots, by assuming that Eve observes $\mathbf{Y}_2$ and $\mathbf{h}_2$, and is given (by a genie) the side information $K$, $\breve{\mathbf{W}}^{[\mathcal{U}_1]}$ and $\mathbf{h}_1$. Note that $\mathbf{X}=\mathbf{X}^{[\mathcal{D}_1]}\cup\mathbf{X}^{[\mathcal{U}_1]}$, where $\mathbf{X}^{[\mathcal{U}_1]}$ plays the role of interference and is not used in the key generation. To bound $E_2$, however, we need Eve to decode the interference given the genie-aided side information.

Given $\mathbf{h}_1$ and $\mathbf{h}_2$, Eve is able to partition $\mathbf{X}$ as

$$\mathbf{X}=\mathbf{X}_{\mathcal{M}^+}\cup\mathbf{X}_{\mathcal{M}^-}, \quad (58)$$

where $\mathcal{M}^+=\{m|m=1,\ldots,M,\text{ and } h_{1m}\ge h_{2m}\}$, $\mathcal{M}^-=\{1,\ldots,M\}/\mathcal{M}^+$, $\mathbf{X}_{\mathcal{M}^+}=\{\mathbf{X}_m|m\in\mathcal{M}^+\}$, and $\mathbf{X}_{\mathcal{M}^-}=\{\mathbf{X}_m|m\in\mathcal{M}^-\}$. We consider the decoding of $\mathbf{X}_{\mathcal{M}^+}$ and $\mathbf{X}_{\mathcal{M}^-}$ separately as in the following subsections.

### A. Decoding of $\mathbf{X}_{\mathcal{M}^-}$

We note that $\mathbf{X}_{\mathcal{M}^-}$ can be partitioned as

$$\mathbf{X}_{\mathcal{M}^-}=\mathbf{X}_{\mathcal{M}^-}^{[\mathcal{D}_2]}\cup\mathbf{X}_{\mathcal{M}^-}^{[\mathcal{U}_2]}. \quad (59)$$

Based on $\mathbf{Y}_{2m}$ and side information $\check{W}_m^{[\mathcal{U}_1]}$, Eve performs the decoding of $\mathbf{X}_{\mathcal{M}^-}$ for each time slot $m\notin\mathcal{M}^+$ independently. The decoding is performed in two steps:

*1) Decoding of $\mathbf{X}_{\mathcal{M}^-}^{[\mathcal{D}_2]}$:* For each $m\in\mathcal{M}^-$, Eve decodes $\mathbf{X}_m^{[\mathcal{D}_2]}$ (decodable layers for Eve) directly based on $\mathbf{Y}_{2m}$ without using side information.

*2) Decoding of $\mathbf{X}_{\mathcal{M}^-}^{[\mathcal{U}_2]}$:* After subtracting $\mathbf{X}_m^{[\mathcal{D}_2]}$ decoded previously, Eve attempts the decoding of $\mathbf{X}_m^{[\mathcal{U}_2]}$ using the side information $\check{W}_m^{[\mathcal{U}_1]}$. More specifically, considering the decoding of $\mathbf{X}_m^{[l]}$ for layer $l\in\mathcal{U}_{2m}$, we use $\check{W}_m^{[l]}$, which is available since we have $\mathcal{U}_{2m}\subset\mathcal{U}_{1m}$ and therefore $\check{W}_m^{[l]}\in\check{W}_m^{[\mathcal{U}_1]}$. We denote by $\mathcal{C}^{[l]}(\check{W}_m^{[l]})$ the thinned codebook corresponding to the genie-informed message $\check{W}_m^{[l]}$. The size of $\mathcal{C}^{[l]}(\check{W}_m^{[l]})$ is $2^{N\hat{r}_m^{[l]}}$, where $\hat{r}_m^{[l]}$ is given by (43). Eve attempts to decode $\mathbf{X}_m^{[l]}$ using $\mathcal{C}^{[l]}(\check{W}_m^{[l]})$ after subtracting the layers lower than $l$, denoted by $\mathbf{X}_m^{[1:(l-1)]}$.

For any typical sequences $\mathbf{X}_m^{[l]}$ and $\mathbf{Y}_{2m}$, it can be shown that

$$I\left(\mathbf{X}_m^{[l]};\mathbf{Y}_{2m}|\mathbf{X}_m^{[1:(l-1)]}\right) \geq N\left[\log\left(1+\frac{h_{2m}p^{[l]}}{1+h_{2m}\sum_{i=l+1}^{L}p^{[i]}}\right)-\epsilon\right].$$

Hence, Eve is able to decode $\mathbf{X}_m^{[l]}$ with an arbitrarily small error probability when $N\to\infty$. By performing decoding for all $l\in\mathcal{U}_{2m}$ successively, Eve decodes $\mathbf{X}_m^{[\mathcal{U}_2]}$.

### B. Decoding of $\mathbf{X}_{\mathcal{M}^+}$

We note that $\mathbf{X}_{\mathcal{M}^+}$ can be partitioned as

$$\mathbf{X}_{\mathcal{M}^+}=\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_2]}\cup\left(\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]}\cap\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}\right)\cup\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_1]}, \tag{60}$$

and Eve performs the decoding of $\mathbf{X}_{\mathcal{M}^+}$ through the following three steps:

*1) Decoding of* $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_2]}$*:* Eve decodes $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_2]}$ directly based on $\mathbf{Y}_2$ without using side information.

*2) Decoding of* $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]}\cap\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$*:* Eve decodes $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]}\cap\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$ jointly based on a list decoding argument, which is explained in details as in the following. A similar argument based on list decoding was given in [12].

**Definition 3.** Sequence $\underline{\mathbf{X}}_m$ is the concatenation of the codewords sent from the group of communication codebooks $\mathcal{C}^{[\mathcal{D}_{1m}\cap\mathcal{U}_{2m}]}$ (i.e. $\underline{\mathbf{X}}_m=\left[\mathbf{X}_m^{[l_{2m}+1]},\ldots,\mathbf{X}_m^{[l_{1m}]}\right]$). The concatenation of sequences $\underline{\mathbf{X}}_m$ for all $m\in\mathcal{M}^+$ is called a super-sequence, denoted by $\underline{\mathbf{X}}$.

The length of sequence $\underline{\mathbf{X}}_m$ is $N(l_{1m}-l_{2m})$, and the length of super-sequence $\underline{\mathbf{X}}$ is therefore $N\sum_{m\in\mathcal{M}^+}(l_{1m}-l_{2m})$. Therefore, the length of a super-sequence depends on the channel realizations of $\mathbf{h}_1$ and $\mathbf{h}_2$ for a finite $M$. However, as $M\to\infty$, it can be seen that the length does not depend on the channel realizations.

As shown in Fig. 8, Eve generates two lists of such super-sequences $\mathcal{L}$ and $\mathcal{T}$ based on genie-provided secret key $K$ and joint-typicality, respectively.

First, given a secret key $K$, Eve narrows down to bin $\mathcal{B}(K)$ in the key generation codebook. Since the mapping function $g$ is deterministic (one-to-one) and encoding in the communication phase is also deterministic, Eve is able to generate $\mathcal{L}(K)$, a list of super-sequences each of which corresponds to a codeword in bin $\mathcal{B}(K)$. Hence, the size of $\mathcal{L}(K)$ is $\|\mathcal{L}(K)\|=2^{nR_s}$.

For each $m\in\mathcal{M}^+$ and any possible sequence $\underline{\mathbf{X}}_m$, we define that

- if $(\mathbf{X}_m^{[\mathcal{D}_1\cap\mathcal{U}_2]},\mathbf{Y}_{2m})$ are jointly typical when $\mathbf{X}_m^{[\mathcal{D}_2]}$ are decoded and substraced from $\mathbf{Y}_{2m}$,
$$\gamma(\underline{\mathbf{X}}_m,\mathbf{Y}_{2m})=1;$$
- otherise, $\gamma(\underline{\mathbf{X}}_m,\mathbf{Y}_{2m})=0$.

Eve constructs a list $\mathcal{L}_m$ such that

$$\mathcal{T}_m=\{\underline{\mathbf{X}}_m|\gamma(\underline{\mathbf{X}}_m,\mathbf{Y}_{2m})=1\}. \tag{61}$$

That is, $\mathcal{T}_m$ consists of the sequences such that the corresponding codewords coming from codebooks $\mathcal{C}^{[\mathcal{D}_{1m}\cap\mathcal{U}_{2m}]}$ are jointly typical with $\mathbf{Y}_{2m}$ given that $\mathbf{X}_m^{[\mathcal{D}_2]}$ has already been decoded and canceled. Finally, Eve constructs a list $\mathcal{T}$ by concatenating sequences in $\mathcal{T}_m$ for all $m\in\mathcal{M}^+$.

Suppose that $\underline{\mathbf{X}}$ is the super-sequence corresponding to the transmitted codewords $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]}\cap\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$. Given the two lists $\mathcal{L}(K)$ and $\mathcal{T}$, Eve attempts to find $\underline{\mathbf{X}}$. Eve declares that $\underline{\mathbf{X}}$ were sent, if $\underline{\mathbf{X}}$ is the only common super-sequence in both $\mathcal{L}(K)$ and $\mathcal{T}$. She declares an error if there is no super-sequence or more than one super-sequences in $\mathcal{L}(K)\cap\mathcal{T}$. Hence, there are two error events correspondingly,

$\mathcal{E}_1$ : $\underline{\mathbf{X}}\notin\mathcal{L}(K)\cap\mathcal{T}$,
$\mathcal{E}_2$ : there exists $\tilde{\underline{\mathbf{X}}}\neq\underline{\mathbf{X}}$, and $\tilde{\underline{\mathbf{X}}}\in\mathcal{L}(K)\cap\mathcal{T}$.

The Asymptotic Equipartition Property (AEP) implies that $\Pr(\mathcal{E}_1)\leq\epsilon_1$, where $\epsilon_1\to0$ as $n\to\infty$. $\Pr(\mathcal{E}_2)$ is bounded as the follows:

$$\begin{aligned}\Pr(\mathcal{E}_2)&\leq\mathbb{E}\left\{\sum_{\tilde{\underline{\mathbf{X}}}\in\mathcal{T},\tilde{\underline{\mathbf{X}}}\neq\underline{\mathbf{X}}}\Pr\left(\tilde{\underline{\mathbf{X}}}\in\mathcal{L}(K)\right)\right\}\\&\leq\mathbb{E}\left\{\|\mathcal{L}\|2^{-nR_s}\right\},\end{aligned} \tag{62}$$

where $\|\mathcal{L}\|$ represents the size of the list $\mathcal{L}$, and (62) follows from the uniform distribution of super-sequences in $\mathcal{L}(K)$.

To proceed, we need to give a bound on $\|\mathcal{L}\|$. We denote the size of $\mathcal{L}_m$ to be $\|\mathcal{L}_m\|$. For any $m\in\mathcal{M}^+$, $\|\mathcal{L}_m\|$ can be bounded as the follows:

$$\begin{aligned}\|\mathcal{L}_m\|&=\mathbb{E}\left\{\sum_{\underline{\mathbf{X}}_m}\gamma(\underline{\mathbf{X}}_m,\mathbf{Y}_{2m})\right\}\\&\leq1+\sum_{\tilde{\underline{\mathbf{X}}}_m\neq\underline{\mathbf{X}}_m}\mathbb{E}\left\{\gamma(\underline{\mathbf{X}}_m,\mathbf{Y}_{2m})\right\},\\&\leq1+2^{N\left(\sum_{l=l_{2m}+1}^{l_{1m}}r^{[l]}\right)}2^{N\left[-\log\left(1+\frac{h_{2m}\sum_{l=l_{2m}+1}^{l_{1m}}p^{[l]}}{1+h_{2m}\sum_{l=l_{1m}+1}^{L}p^{[l]}}\right)+\epsilon_2\right]}\\&\leq2^{N\left\{\sum_{l=l_{2m}+1}^{l_{1m}}\left[r^{[l]}-\log\left(1+\frac{h_{2m}p^{[l]}}{1+h_{2m}\sum_{i=l+1}^{L}p^{[i]}}\right)\right]+\epsilon_3\right\}},\end{aligned}$$

where $\epsilon_2,\epsilon_3\to0$ as $N\to\infty$. The size of $\mathcal{L}$ is then bounded as

$$\begin{aligned}\|\mathcal{L}\|&=\prod_{m\in\mathcal{M}^+}\|\mathcal{L}_m\|\\&\leq2^{N\sum_{m\in\mathcal{M}^+}\left\{\sum_{l=l_{2m}+1}^{l_{1m}}\left[r^{[l]}-\log\left(1+\frac{h_{2m}p^{[l]}}{1+h_{2m}\sum_{i=l+1}^{L}p^{[i]}}\right)\right]+\epsilon_3\right\}}.\end{aligned}$$

As $M\to\infty$, by following steps similar as those for deriving (57), we have

$$\|\mathcal{L}\|\leq2^{n(R_s-\epsilon_4)}, \tag{63}$$

Now we can combine (62) and (63) to obtain that

$$\Pr(\mathcal{E}_2)\leq2^{-n\epsilon_4}\to0, \tag{64}$$

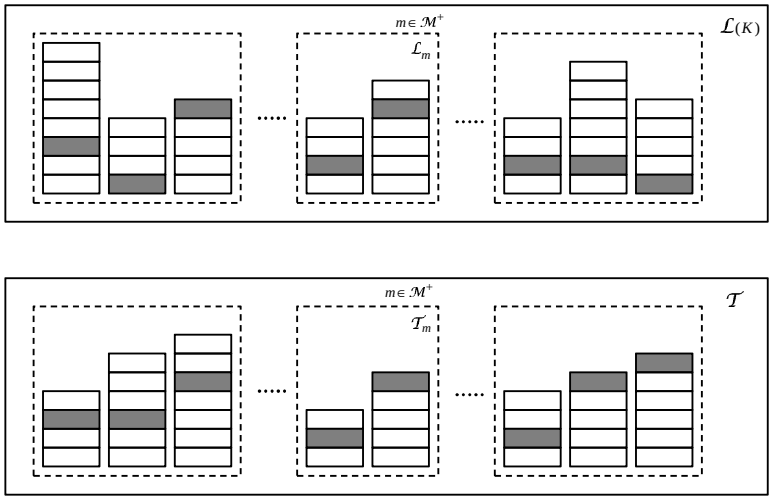


Fig. 8. Two lists of super-sequences: (a) list $\mathcal{L}(K)$ constructed based on genie-provided $K$, (b) list $\mathcal{T}$ constructed based on joint-typicality.

as $n \to \infty$. Hence, the average error probability for decoding $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]} \cap \mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$ is bounded by

$$\Pr(\mathcal{E}_1 \cup \mathcal{E}_2) \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2) \to 0,$$

as $n \to \infty$. Thus, Eve is able to find the right super-sequence $\underline{\underline{\mathbf{X}}}$ with a vanishing error probability. Since $\underline{\underline{\mathbf{X}}}$ and the group of codewords $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]} \cap \mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$ are related by a one-to-one mapping, we conclude that Eve is able to decode $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]} \cap \mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$ with a vanishing error probability.

*3) Decoding of $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_1]}$:* Eve subtracts $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_2]}$ and $\mathbf{X}_{\mathcal{M}^+}^{[\mathcal{D}_1]} \cap \mathbf{X}_{\mathcal{M}^+}^{[\mathcal{U}_2]}$ from $\mathbf{Y}_2$ based on the two previous decoding procedures, and tries to decode $\mathbf{X}_m^{[\mathcal{U}_1]}$ using the thinned codebooks $\mathcal{C}^{[\mathcal{U}_1]}(\breve{\mathbf{W}}^{[\mathcal{U}_1]})$. The decoding procedure is similar to that discussed in subsection A.2.

Finally, we conclude that Eve is able to decode $\mathbf{X}$ given $\mathbf{Y}_2$, the genie-informed (secret-key) information $K$, and the side information $\breve{\mathbf{W}}^{[\mathcal{U}_1]}$. Hence, Fano's inequality implies that

$$E_2 = H(\mathbf{X}|\mathbf{Y}_2, W, \breve{\mathbf{W}}^{[\mathcal{U}_1]}, \mathbf{h}_1, \mathbf{h}_2) \leq n\delta_n \to 0, \qquad (65)$$

as $n \to \infty$. We thus complete the proof of (48).

## Appendix E
## Proof of Lemma 2

We can rewrite the secrecy key rate $R_s$ as

$$\begin{aligned} R_s &= T_1 - T_2 \\ &= \int_0^\infty T_{1i}(h_1) d[1 - F_1(h_1)] - \int_0^\infty T_{2i}(h_1) d[1 - F_1(h_1)], \end{aligned} \qquad (66)$$

where

$$T_{1i}(h_1) = \int_0^{h_1} \left[ \int_{h_2}^{h_1} \frac{s\rho(s)ds}{1 + sI(s)} \right] d[1 - F_2(h_2)], \qquad (67)$$

and

$$T_{2i}(h_1) = \int_0^{h_1} \left[ \int_{h_2}^{h_1} \frac{h_2\rho(s)ds}{1 + h_2 I(s)} \right] d[1 - F_2(h_2)]. \qquad (68)$$

### A. Evaluation of $T_1$

$T_{1i}(h_1)$ can be evaluated by integrating by part. We have

$$\begin{aligned} T_{1i}(h_1) &= \left[ \int_{h_2}^{h_1} \frac{s\rho(s)ds}{1 + sI(s)} \right] [1 - F_2(h_2)] \Big|_0^{h_1} \\ &\quad - \int_0^{h_1} [1 - F_2(h_2)] \, d \left[ \int_{h_2}^{h_1} \frac{u\rho(u)du}{1 + uI(u)} \right] \\ &= - \int_0^{h_1} \frac{s\rho(s)ds}{1 + sI(s)} + \int_0^{h_1} [1 - F_2(s)] \frac{s\rho(s)ds}{1 + sI(s)} \\ &= - \int_0^{h_1} F_2(s) \frac{s\rho(s)ds}{1 + sI(s)} \end{aligned} \qquad (69)$$

By another integrating by part, we obtain

$$\begin{aligned} T_1 &= \int_0^\infty T_{1i}(h_1) d \left[ 1 - F_1(h_1) \right] \\ &= T_{1i}(h_1) \left[ 1 - F_1(h_1) \right] \Big|_0^\infty - \int_0^\infty \left[ 1 - F_1(h_1) \right] d \left[ T_{1i}(h_1) \right] \\ &= - \int_0^\infty \left[ 1 - F_1(h_1) \right] d \left[ T_{1i}(h_1) \right] \\ &= \int_0^\infty \left[ 1 - F_1(s) \right] F_2(s) \frac{s\rho(s)ds}{1 + sI(s)}. \end{aligned} \qquad (70)$$

### B. Evaluation of $T_2$

$T_{2i}(h_1)$ can be rewritten as

$$\begin{aligned} T_{2i}(h_1) &= \left[\int_{h_2}^{h_1} \frac{h_2\rho(s)ds}{1+h_2I(s)}\right][1-F_2(h_2)]\Big|_0^{h_1} \\ &\quad - \int_0^{h_1}[1-F_2(h_2)]\,d\left[\int_{h_2}^{h_1}\frac{h_2\rho(s)ds}{1+h_2I(s)}\right] \\ &= -\int_0^{h_1}[1-F_2(s)]\,d\left[\int_{h_2}^{h_1}\frac{h_2\rho(s)ds}{1+h_2I(s)}\right]. \end{aligned}$$

Notice that

$$\int_{h_2}^{h_1}\frac{h_2\rho(s)ds}{1+h_2I(s)} = \log\left(1+h_2I(h_2)\right) - \log\left(1+h_2I(h_1)\right),$$

and therefore

$$\frac{d}{dh_2}\left[\int_{h_2}^{h_1}\frac{h_2\rho(s)ds}{1+h_2I(s)}\right] = \frac{I(h_2)-h_2\rho(h_2)}{1+h_2I(h_2)} - \frac{I(h_1)}{1+h_2I(h_1)}.$$

Hence, $T_{2i}(h_1)$ can be written as

$$\begin{aligned} T_{2i}(h_1) = &-\int_0^{h_1}[1-F_2(h_2)]\times \\ &\left[\frac{I(h_2)-h_2\rho(h_2)}{1+h_2I(h_2)} - \frac{I(h_1)}{1+h_2I(h_1)}\right]dh_2. \end{aligned} \tag{71}$$

Furthermore, we have

$$\begin{aligned} \frac{d}{dh_1}T_{2i}(h_1) = &-[1-F_2(h_1)]\left[\frac{I(h_1)-h_1\rho(h_1)}{1+h_1I(h_1)}\right] \\ &+\frac{d}{dh_1}\left[\int_0^{h_1}[1-F_2(h_2)]\frac{I(h_1)}{1+h_2I(h_1)}dh_2\right]. \end{aligned} \tag{72}$$

To proceed, we need to interchange the operation of differentiation with respect to $h_1$ with the operation of integration over $h_2$, where the integral domain is also a function of $h_1$. We use the property that for any real differentiable function $p(x,y)$, we can write

$$\frac{d}{dx}\int_0^x p(x,y)dy = p(x,x) + \int_0^x \frac{\partial p(x,y)}{\partial x}dy. \tag{73}$$

In particular, we have

$$\begin{aligned} &\frac{d}{dh_1}\left[\int_0^{h_1}[1-F_2(h_2)]\frac{I(h_1)}{1+h_2I(h_1)}dh_2\right] \\ &= [1-F_2(h_1)]\frac{I(h_1)}{1+h_1I(h_1)} \\ &\quad + \int_0^{h_1}[1-F_2(h_2)]\frac{\partial}{\partial h_1}\left[\frac{I(h_1)}{1+h_2I(h_1)}\right]dh_2 \\ &= [1-F_2(h_1)]\frac{I(h_1)}{1+h_1I(h_1)} \\ &\quad + \frac{\rho(h_1)}{I(h_1)}\int_0^{h_1}[1-F_2(h_2)]\,d\frac{1}{1+h_2I(h_1)} \\ &= [1-F_2(h_1)]\frac{I(h_1)}{1+h_1I(h_1)} \\ &\quad + \frac{\rho(h_1)}{I(h_1)}\left[\frac{1-F_2(h_1)}{1+h_1I(h_1)} - 1 + \int_0^{h_1}\frac{f_2(h_2)dh_2}{1+h_2I(h_1)}\right], \end{aligned} \tag{74}$$

where we have used integrating by part to get to the last equality.

Putting (74) into (72), we have

$$\frac{d}{dh_1}T_{2i}(h_1) = -\frac{F_2(h_1)\rho(h_1)}{I(h_1)} + \frac{\rho(h_1)}{I(h_1)}\int_0^{h_1}\frac{f_2(h_2)dh_2}{1+h_2I(h_1)}. \tag{75}$$

Now, we can evaluate $T_2$ by

$$\begin{aligned} T_2 &= \int_0^\infty T_{2i}(h_1)d\left[1-F_1(h_1)\right] \\ &= -\int_0^\infty [1-F_1(h_1)]\,dT_{2i}(h_1) \\ &= \int_0^\infty \frac{[1-F_1(h_1)]\,F_2(h_1)\rho(h_1)}{I(h_1)}dh_1 \\ &\quad - \int_0^\infty \frac{[1-F_1(h_1)]\,\rho(h_1)}{I(h_1)}\left[\int_0^{h_1}\frac{f_2(h_2)dh_2}{1+h_2I(h_1)}\right]dh_1. \end{aligned} \tag{76}$$

### C. Evaluation of $R_s = T_1 - T_2$

Using (70) and (76), and replacing the variable $h_1$ and $h_2$ with $x$ and $y$, respectively, we have

$$\begin{aligned} R_s &= \int_0^\infty \frac{[1-F_1(x)]\,\rho(x)}{I(x)}\left[\int_0^x\frac{f_2(y)dy}{1+yI(x)} - \frac{F_2(x)}{1+xI(x)}\right] \\ &= \int_0^\infty [1-F_1(x)]\,\rho(x)\left[\int_0^x\frac{F_2(y)dy}{[1+yI(x)]^2}\right], \end{aligned} \tag{77}$$

which is (27).

## Appendix F
## Proof of Lemma 2

The functional of (27) is defined by

$$L\left(x, I(x), I'(x)\right) = -[1-F_1(x)]\,I'(x)\left[\int_0^x\frac{F_2(y)dy}{[1+yI(x)]^2}\right].$$

A necessary condition for a maximum of the integral of $L(x, I(x), I'(x))$ over $x$ is a zero variation of the functional. For characterizing the optimal $I(x)$, the Eüler-Lagrangian equation [32] gives a necessary condition denoted by

$$\frac{\partial L}{\partial I} - \frac{d}{dx}\left(\frac{\partial L}{\partial I'}\right) = 0, \tag{78}$$

for which we have,

$$\frac{\partial L}{\partial I} = 2\,[1-F_1(x)]\,I'(x)\int_0^x\frac{yF_2(y)dy}{[1+yI(x)]^3}, \tag{79}$$

$$\frac{\partial L}{\partial I'} = -[1-F_1(x)]\int_0^x\frac{F_2(y)dy}{[1+yI(x)]^2}, \tag{80}$$

$$\begin{aligned} \frac{d}{dx}\frac{\partial L}{\partial I'} &= f_1(x)\int_0^x\frac{F_2(y)dy}{[1+yI(x)]^2} \\ &\quad - [1-F_1(x)]\frac{d}{dx}\int_0^x\frac{F_2(y)dy}{[1+yI(x)]^2}, \end{aligned} \tag{81}$$

with

$$\frac{d}{dx}\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2} = \frac{F_2(x)}{[1+xI(x)]^2} - 2I'(x)\int_0^x \frac{yF_2(y)dy}{[1+yI(x)]^3}. \quad (82)$$

Using (79), (81), and (82) in (78), we have

$$\int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2} = \frac{[1-F_1(x)]\,F_2(x)}{[1+xI(x)]^2 f_1(x)}. \quad (83)$$

Hence, we proved Lemma 2.

## Appendix G
## Proof of (35)

According to Lemma 1, the secrecy rate is

$$\begin{aligned} R_s^{[1]} &= \Pr\left[R^{[1]} \le \log(1+h_1P)\right]\mathbb{E}_{h_2}\left[R^{[1]} - \log(1+h_2P)\right]^+ \\ &= \Pr\{h_1 \ge h_1^*\}\int_0^{h_1^*}\left[R^{[1]} - \log(1+h_2P)\right] f_2(h_2)dh_2 \\ &= \exp\left(-\frac{h_1^*}{\lambda_1}\right)\times \\ &\quad \left[R^{[1]}F_2(h_1^*) - \int_0^{h_1^*}\log(1+h_2P)f_2(h_2)dh_2\right], \end{aligned} \quad (84)$$

where $h_1^* = \left[\exp(R^{[1]})-1\right]/P$. By using integrating by part for the integral in (84), we have

$$\begin{aligned} R_s^{[1]} &= \exp\left(-\frac{h_1^*}{\lambda_1}\right)\int_0^{h_1^*}\frac{[1-\exp(-h_2/\lambda_2)]\,P}{1+h_2P}dh_2 \\ &= \exp\left(-\frac{h_1^*}{\lambda_1}\right)\left[R^{[1]} - \int_0^{h_1^*}\frac{\exp(-h_2/\lambda_2)\,P}{1+h_2P}dh_2\right]. \end{aligned}$$

By letting $t = (1+h_2P)/(\lambda_2 P)$, we have

$$\begin{aligned} R_s^{[1]} &= \exp\left(-\frac{h_1^*}{\lambda_1}\right)\times \\ &\quad \left[R^{[1]} - \exp\left(\frac{1}{\lambda_2 P}\right)\int_{\frac{1}{\lambda_2 P}}^{\frac{1+h_1^*P}{\lambda_2 P}}\frac{\exp(-t)}{t}dt\right] \\ &= \exp\left(-\frac{h_1^*}{\lambda_1}\right)\times \\ &\quad \left\{R^{[1]} - \exp\left(\frac{1}{\lambda_2 P}\right)\left[E_i\left(\frac{1+h_1^*P}{\lambda_2 P}\right) - E_i\left(\frac{1}{\lambda_2 P}\right)\right]\right\} \end{aligned}$$

By using $h_1^* = \left[\exp(R^{[1]})-1\right]/P$, we can obtain (35).

## Appendix H
## Proof of (36)

We can write

$$\begin{aligned} \int_0^x \frac{F_2(y)dy}{[1+yI(x)]^2} &= \int_0^x \frac{1-\exp(-y/\lambda_2)}{[1+yI(x)]^2}dy \\ &= \underbrace{\int_0^x \frac{dy}{[1+yI(x)]^2}}_{T_3} - \underbrace{\int_0^x \frac{\exp(-y/\lambda_2)\,dy}{[1+yI(x)]^2}}_{T_4}, \end{aligned} \quad (85)$$

and evaluate $T_3$ and $T_4$ separately. First, we have

$$T3 = \frac{x}{1+xI(x)}. \quad (86)$$

To evaluate $T_4$, we have

$$\begin{aligned} T_4 &= -\int_0^x \frac{\exp(-y/\lambda_2)}{I(x)}d\left[\frac{1}{1+yI(x)}\right] \\ &= -\frac{1}{I(x)}\frac{\exp(-y/\lambda_2)}{1+yI(x)}\Big|_0^x - \frac{1}{\lambda_2 I(x)}\int_0^x \frac{\exp(-y/\lambda_2)\,dy}{1+yI(x)} \\ &= \frac{1}{I(x)}\left[1 - \frac{\exp(-x/\lambda_2)}{1+xI(x)}\right] \\ &\quad \underbrace{-\frac{1}{\lambda_2 I^2(x)}\int_0^x \frac{\exp(-y/\lambda_2)}{1+yI(x)}d[1+yI(x)]}_{T_5}. \end{aligned} \quad (87)$$

By letting $1+yI(x) = t$, we have

$$\begin{aligned} T_5 &= \frac{\exp(1/\lambda_2 I(x))}{\lambda_2 I^2(x)} \\ &\quad \times \int_1^{1+xI(x)}\frac{\exp(-t/\lambda_2 I(x))}{t/\lambda_2 I(x)}d\left[\frac{t}{\lambda_2 I(x)}\right] \\ &= \frac{1}{\lambda_2 I^2(x)}\exp\left(\frac{1}{\lambda_2 I(x)}\right)E_i\left(\frac{t}{\lambda_2 I(x)}\right)\Big|_{1+xI(x)}^1 \\ &= \frac{1}{\lambda_2 I^2(x)}\exp\left(\frac{1}{\lambda_2 I(x)}\right)\times \\ &\quad \left[E_i\left(\frac{1}{\lambda_2 I(x)}\right) - E_i\left(\frac{1+xI(x)}{\lambda_2 I(x)}\right)\right]. \end{aligned} \quad (88)$$

Combining (86), (87), (88) and (85), we can obtain (36).

## References


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

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

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

[4] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[5] R. Liu, I. Maric, R. D. Yates, and P. Spasojevic, "The discrete memoryless multiple access channel with confidential messages," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, July 2006, pp. 957–961.

[6] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[7] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wire-tap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[8] R. Liu, I. Maric, P. Spasojević, and R. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy capacity regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[9] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai (Shitz), and S. Verdú, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[10] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2008.

[11] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1895–1911, Oct. 1998.

[12] 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.

[13] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[14] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of indepedent parallel channels," in *Proc. Allerton Conference on Commun. Contr. Computing*, Monticello, IL, USA, Sept. 2006.

[15] X. Tang, R. Liu, P. Spasojević, and H. V. Poor, "On the throughput of secure hybrid-ARQ protocols for Gaussian block-fading channels," *IEEE Trans. on Inf. Theory*, vol. 55, no. 4, pp. 1575–1591, Apr. 2009.

[16] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, 1993.

[17] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[18] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. on Inf. Theory*, vol. 46, no. 2, pp. 344–366, Feb. 2000.

[19] ——, "Secrecy capacities for multiple terminals," *IEEE Trans. on Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[20] ——, "Secrecy capacities for multiterminal channel models," *IEEE Trans. on Inf. Theory*, vol. 54, no. 6, p. 24372452, Jun. 2008.

[21] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels - part I: Definitions and a completeness result," *IEEE Trans. on Inf. Theory*, vol. 49, no. 4, pp. 822–831, Apr. 2003.

[22] ——, "Secret-key agreement over unauthenticated public channels - part II: The simulatability condition," *IEEE Trans. on Inf. Theory*, vol. 49, no. 4, pp. 832–838, Apr. 2003.

[23] ——, "Secret-key agreement over unauthenticated public channels - part III: Privacy amplification," *IEEE Trans. on Inf. Theory*, vol. 49, no. 4, pp. 839–851, Apr. 2003.

[24] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret key generation using correlated sources and noisy channels," in *Proc. IEEE Int. Symp. Inf. Thoery*, Toronto, Canada, Jul. 2008, pp. 1005–1009.

[25] V. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels: A secret key - secret message rate tradeoff region," in *Proc. IEEE Int. Symp. Inf. Thoery*, Toronto, Canada, Jul. 2008, pp. 1010–1014.

[26] X. Tang, R. Liu, P. Spasojević, and H. V. Poor, "The Gaussian wiretap channel with a helping interferer," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, Canada, Jul. 2008 2008.

[27] ——, "Interference assisted secret communication," *IEEE Trans. on Inf. Theory*, to appear.

[28] S. Shamai (Shitz), "A broadcast strategy for the Gaussian slowly fading channel," in *Proc. IEEE Int. Symp. Inf. Thoery*, Ulm, Germany, Jun. 29 - Jul. 4 1997.

[29] S. Shamai (Shitz) and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. on Inf. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[30] M. Abdel Latif, A. Sultan, and H. El Gamal, "ARQ based secret key sharing," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Dresden, Germany, June 14-18 2009.

[31] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[32] I. Gelfand and S. Fomin, *Calculus of Variations*. Englewood Cliffs, NJ: Prentice-Hall, 1963.