On the Fast Fading Gaussian Wiretap Channel With Statistical Channel State Information at the Transmitter

Pin-Hsun Lin, *Member, IEEE*, and Eduard Jorswieck, *Senior Member, IEEE*

***Abstract*— In this paper, we investigate the ergodic secrecy capacity of the fast fading Gaussian wiretap channel when only the statistics of the channel state information are known at the transmitter. We derive conditions for the existence of degradedness and a positive ergodic secrecy capacity under the usual stochastic order, the convex order, and the increasing convex order between the legitimate and eavesdropper channels. For more general orders, we prove the secrecy capacity of layered erasure wiretap channels and propose a layered signaling for the achievable scheme, and we derive an upper bound on the capacity for fast fading Gaussian wiretap channels. Finally, the numerical results show that under Nakagami-*m* fast fading channels, the proposed layered signaling outperforms the Gaussian codebook in several cases. In particular, in certain cases, the Gaussian codebook can achieve only a zero secrecy rate, whereas the proposed scheme achieves positive secrecy rates. Therefore, the connectivity of wireless networks can be significantly improved by the proposed scheme.**

***Index Terms*— Information theoretic security, degraded wiretap channel, fast-fading, stochastic order, ergodic secrecy capacity.**

I. Introduction

KEY-BASED enciphering is traditionally used to ensure the security of data transmission. However, key distribution and management may be challenging tasks [1] for secure wireless communications because of the additional control signaling and feedback channel management. By contrast, the physical-layer security introduced in [2] and [3] is appealing because of its keyless nature. One of the major difficulties regarding physical-layer security is the characterization of the secrecy capacities of wiretap channels. The secrecy capacity is the maximum achievable secrecy rate between a transmitter and a legitimate receiver when a weak secrecy constraint is imposed to prevent information from being available to an eavesdropper [2], [3]. Further enhancements can be attained by employing multiple antennas, e.g., as in [4]–[7]. In wireless environments where each node has a single antenna, the time-varying characteristics of fading channels can also be exploited to enhance the secrecy capacity [8]. However, to derive the optimal input distribution of the secrecy capacity, at least perfect knowledge of the legitimate channel state information at the transmitter (CSIT) is required by the schemes proposed in [5]–[8]. Because of limited feedback bandwidth and the delay caused by channel estimation, the transmitter may not be able to track channel realizations instantaneously if they vary rapidly. Thus, for fast fading channels, it is more practical to consider the case with only partial CSIT of the legitimate channel [9]–[14]. In this case, when the transmitter has multiple antennas, the problem of determining the ergodic secrecy capacity is solved only for the special case in which both Bob's and Eve's* channels are Rayleigh distributed (but with different statistics) and the entries of each channel vector are identically and independently distributed (i.i.d.) [12]. This result can be further extended to cases in which both Bob and Eve have multiple antennas with total or per-antenna power constraints [15]. For the extreme case in which the channel gains are unknown to all parties, with Bob's and Eve's channels being fast Rayleigh faded, the secrecy capacity can be determined for a discrete channel input [16]. However, for more general settings, e.g., Bob's and Eve's channels do not belong to the same type of distribution, only some lower and upper bounds on the secrecy capacity are known [10], [13], [17], [18]. More specifically, although the general secrecy capacity formula has been presented in [3], the optimal means of selection of the auxiliary random variable and channel prefixing remain unknown for partial CSIT cases in general.

Channel enhancement [19] is a powerful tool for transforming non-degraded channels, such as multiple-input multiple-output (MIMO) channels into degraded ones, such that for several classes of channels, the design of auxiliary random variables and channel prefixing can be avoided. The channel enhancement technique was originally developed to prove the capacity of the MIMO Gaussian broadcast channel [19]. It was later applied to prove the capacity of the MIMO Gaussian wiretap channel with a covariance matrix constraint [7], [20] and endows the proof with an intuitive explanation. However, traditional channel enhancement relies on full knowledge of

Manuscript received September 4, 2014; revised March 31, 2015 and August 19, 2015; accepted August 25, 2015. Date of publication September 3, 2015; date of current version October 30, 2015. This work was performed in the framework of the European research project DIWINE, which was supported in part by the European Union under its FP7 ICT Objective 1.1-The Network of the Future. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shantanu D. Rane.

The authors are with the Communications Laboratory, Electrical and Computer Engineering Department, Technische Universität Dresden, Dresden 01062, Germany (e-mail: pin-hsun.lin@tu-dresden.de; eduard.jorswieck@tu-dresden.de).

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.2015.2476464

*In the following, for simplicity and clarity, we use Alice, Bob and Eve to represent the transmitter, the legitimate receiver, and the eavesdropper, respectively.

the CSIT, and it is not clear how to apply the concept of enhancement to a fast fading channel with partial CSIT.

In this paper, we investigate the relation between different stochastic orders and the ergodic secrecy capacity of a wiretap channel under fast fading. More specifically, even without perfect CSIT, we may still be able to *stochastically* compare the *channel orders*. Each node is assumed to be equipped with a single antenna, and only statistical CSIT is available for both channels. The main contributions of this paper are as follows.

- We characterize the relation between different stochastic orders and the existence of the equivalent degraded wiretap channel with a positive ergodic secrecy capacity. The considered stochastic orders include the usual stochastic order, the convex order, and the increasing convex order. Several examples with practical channel distributions are also illustrated. Note that such a characterization of fast fading wiretap channels with partial CSIT is lacking in the literature.
- We derive the upper bound on the capacity of the fast fading wiretap channel with statistical CSIT by applying the *channel enhancement* technique [21].
- We propose a layered signaling based on the deterministic scheme [21], [22] to derive the lower bound on the ergodic secrecy capacity. When the wiretap channel is not degraded, the layered scheme enables transmissions only on those layers where Bob's channels are superior to Eve's, which cannot be achieved using the traditional Gaussian codebook.
- Using several numerical examples, we show that under Nakagami-m fading channels, with only three binary layers, the proposed layered scheme can outperform the Gaussian codebook.

Notation: In this paper, upper-case normal letters denote random variables. The mutual information between two random variables is denoted by $I(.;.)$. The complementary cumulative distribution function (CCDF) is denoted by $\bar{F}_X(x) = P(X \geq x)$, $F_X(x)$ is the CDF of X. The probability mass function (PMF) and probability density function (PDF) are denoted by P and f, respectively. $E[.]$ denotes the expectation; $H(.)$ and $h(.)$ denote the entropy and differential entropy, respectively. $\text{supp}(f) = \{x \in \mathcal{X} | f(x) \neq 0\}$, where $f : \mathcal{X} \to \mathbb{R}$. $Unif(a, b)$ denotes the uniform distribution between a and b. The logarithms used in the paper are all of base 2.

The remainder of the paper is organized as follows. In Section II, we introduce the system model under consideration. In Section III, we discuss our first main result, which partially characterizes the relation between the existence of a degraded wiretap channel with a positive secrecy capacity and several stochastic orders among Bob's and Eve's channels. In Section IV-A, we prove that for layered erasure wiretap channels, the binary layered scheme can be used to determine the secrecy capacity with only statistical CSIT for both channels. The secrecy capacity lower and upper bounds with examples are provided in Section IV-B. In the same section, we also derive both lower and upper bounds on the capacity for the fast fading Gaussian wiretap channel. In Section V, a numerical simulation is presented to illustrate our results. Finally, Section VI concludes the paper.

II. System Model

The fast fading wiretap channel under consideration [3] is

$$Y_r = \sqrt{H_r} X + Z_r, \tag{1}$$
$$Y_e = \sqrt{H_e} X + Z_e, \tag{2}$$

where H_r and H_e are real-valued non-negative independent random variables denoting the squares of Bob's and Eve's fading channels, with CCDFs of $\bar{F}_{H_r}$ and $\bar{F}_{H_e}$, respectively. The channel input is denoted by X. Without loss of generality, we normalize the channel input power constraint to $E[X^2] \leq 1$. Note that cases of realizations of Bob's and Eve's channels with negative values can be modeled by an additional phase rotation, which can be de-rotated at the receiver because of the full CSI assumption at the receiver. Because this phase rotation is independent of other variables, we can construct the equivalent channel model as in (1) and (2) without affecting the capacity. We assume that Alice knows only the distributions but not the instantaneous realizations of H_r and H_e. The noise variables Z_r and Z_e at Bob and Eve, respectively, are independent additive white Gaussian noise (AWGN) with zero mean and unit variance.

From [3], under the assumptions that Alice does not know the instantaneous h_r and h_e, Bob knows only h_r, and Eve knows both h_r and h_e, where H_r is independent of H_e, we know that the secrecy capacity of this channel can be represented by

$$\begin{aligned} C_s &\overset{(a)}{=} \max_{p(x|u), p(u)} I(U; Y_r, H_r) - I(U; Y_e, H_e, H_r) \\ &\overset{(b)}{=} \max_{p(x|u), p(u)} I(U; Y_r|H_r) - I(U; Y_e|H_e), \end{aligned} \tag{3}$$

where in (a), we start from the result for the discrete memoryless wiretap channel [3] and treat H_r and $\{H_e, H_r\}$ as channel outputs [23] at Bob and Eve, respectively, whereas in (b), we use the chain rule of mutual information and the fact that H_r and H_e are independent of X and U. Finally, after applying the quantization scheme used in [24, Proof of Theorem 3.3 and Remark 3.8], the capacity of the Gaussian wiretap channel can be derived.

III. Degradedness and Stochastic Orders

In this section, we discuss the relation between the existence of an equivalent degraded wiretap channel and different stochastic orders between Bob's and Eve's channels. This relation is helpful for distinguishing the existence of a positive ergodic secrecy capacity. In addition, our proofs of the lower and upper bounds also strongly depend on the characterization of the stochastic orders. We first introduce the following definitions, which are important for deriving the main results of this work.

Definition 1 [1, p. 373]: A wiretap channel is physically degraded if the transition distribution function satisfies $p_{Y_r Y_e|X}(\cdot|\cdot) = p_{Y_r|X}(\cdot|\cdot) p_{Y_e|Y_r}(\cdot|\cdot)$, *i.e., X, Y_r, and Y_e form a Markov chain $X \to Y_r \to Y_e$. A wiretap channel is*

stochastically degraded if its conditional marginal distribution is the same as that of a physically degraded wiretap channel, i.e., there exists a distribution $\tilde{p}_{Y_e|Y_r}(\cdot|\cdot)$ *such that* $p_{Y_e|X}(y_e|x) = \sum_{y_r} p_{Y_r|X}(y_r|x)\tilde{p}_{Y_e|Y_r}(y_e|y_r)$.

Definition 2 [1, Lemma 2.1]: Two wiretap channels have the same capacity equivocation region if they have the same marginal channel transition probability distributions $P_{Y_r|X}(\cdot|\cdot)$ *and* $P_{Y_e|X}(\cdot|\cdot)$.

In the following, we call a stochastically degraded channel simply a degraded channel because the capacities of wiretap channels depend only on their marginal distributions. Based on the definition of degradedness, we present the following definition.

Definition 3: Define the set of pairs of fast fading channels (H_r, H_e) *as*

$$\mathcal{S}_{\mathcal{D}^+} = \{(H_r, H_e) : \textit{the fast fading Gaussian wiretap channel is degraded and } C_s > 0\}.$$

Based on the above definitions, we can derive our first result.

Theorem 1: With statistical CSI for both channels at the transmitter and the power constraint $E[X^2] \leq 1$, *if* $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$, *then the ergodic secrecy capacity is*

$$C_s = \frac{1}{2}\left\{E_{H_r}[\log(1+H_r)] - E_{H_e}[\log(1+H_e)]\right\}. \quad (4)$$

The proof follows the standard one for the case with full CSIT but is reproduced in Appendix I with appropriate modifications for our assumption of only statistical CSIT. In the following, we introduce several stochastic orders.

Definition 4 [25, (1.A.7), (3.A.1), (4.A.1), and (5.A.1)]: For given random variables X and Y, the *usual stochastic order (st)*, the *convex order (cx)*, the *concave order (cv)*, the *increasing convex order (icx)*, the *increasing concave order (icv)*, and the *Laplace transform order (Lt)* are respectively defined as follows:

$X \leq_{st} Y$, if $E[f(X)] \leq E[f(Y)]$ for all increasing f,

$X \leq_{cx(cv)} Y$, if $E[f(X)] \leq E[f(Y)]$ for all convex (concave) f,

$X \leq_{icx(icv)} Y$, if $E[f(X)] \leq E[f(Y)]$ for all increasing convex (concave) f,

$X \leq_{Lt} Y$, if $E[\exp(-sX)] \leq E[\exp(-sY)]$ for all $s > 0$.

Definition 5: Define the following sets: $\mathcal{S}_{st} = \{(H_r, H_e) : H_r \geq_{st} H_e\}$, $\mathcal{S}_{cx} = \{(H_r, H_e) : H_r \geq_{cx} H_e\}$, *and* $\mathcal{S}_{icx} = \{(H_r, H_e) : H_r \geq_{icx} H_e\}$.

Note that the stochastic orders in Definition 4 can be further represented by the following relations, which are more easily evaluated.

Theorem 2 [25, (1.A.3), (3.A.7), and (4.A.5)]: For random variables X *and* Y, $X \leq_{st} Y$ *if and only if* $\bar{F}_X(x) \leq \bar{F}_Y(x)$ *for all* x, *and* $X \leq_{icx} Y$ *if and only if*

$$\int_t^\infty \bar{F}_X(h)dh \leq \int_t^\infty \bar{F}_Y(h)dh \quad (5)$$

for all t. *Moreover,* $X \leq_{cx} Y$ *if and only if* (5) *is valid for all* t *and* $E[X] = E[Y]$. *Similarly,* $X \leq_{icv} Y$ *if and only if*

$$\int_{-\infty}^t \bar{F}_X(h)dh \leq \int_{-\infty}^t \bar{F}_Y(h)dh \quad (6)$$

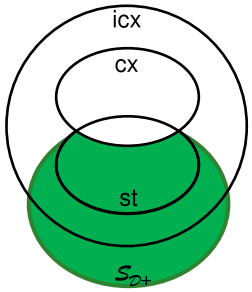

Fig. 1. The relation between the various stochastic orders and the set $\mathcal{S}_{\mathcal{D}^+}$, which is encircled by the green line.

for all t. *Finally,* $X \leq_{cv} Y$ *if and only if* (6) *is valid for all* t *and* $E[X] = E[Y]$.

Note that when X and Y are nonnegative, the condition $E[X] = E[Y]$ can be further expressed as $\int_0^\infty \bar{F}_X(h)dh = \int_0^\infty \bar{F}_Y(h)dh$ [26, (1.12)]. Compared with the original expectation, the integral form of the CCDFs, which unifies the expression of the considered stochastic orders as functions of the CCDFs only, highly simplifies the following derivations.

In the following, we classify the relation between $\mathcal{S}_{\mathcal{D}^+}$ and the different stochastic orders. The results are illustrated in Fig. 1.[†]

Lemma 1: The usual stochastic order $H_r \geq_{st} H_e$ *is sufficient but not necessary to generate an equivalent degraded wiretap channel.*

Proof: We first prove that $H_r \geq_{st} H_e$ is sufficient to generate an equivalent degraded wiretap channel by slightly modifying the proof of [21, Lemma 3] and adapting it to the wiretap channel case. Because the same marginal property holds in a wiretap channel [1, Lemma 2.1], we can generate arbitrary joint distributions of H_r and H_e, which do not alter the capacity as long as the marginal distributions remain unchanged. In the following, we denote the channels equivalent[‡] to H_r and H_e by $\tilde{H}_r$ and $\tilde{H}_e$, respectively. Then, we can form the joint CCDF as follows:

$$\bar{F}_{\tilde{H}_r,\tilde{H}_e}(r, e) = \min\{\bar{F}_{H_r}(r), \bar{F}_{H_e}(e)\}, \quad (7)$$

where $\bar{F}_{X,Y}(x, y) \triangleq P(X \geq x, Y \geq y)$, from which it is clear that the marginal distributions are unchanged, i.e., $\bar{F}_{\tilde{H}_r}(r) = \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r, 0) = \bar{F}_{H_r}(r)$ and $\bar{F}_{\tilde{H}_e}(r) = \bar{F}_{\tilde{H}_r,\tilde{H}_e}(0, e) = \bar{F}_{\tilde{H}_e}(e)$. Thus, by the same marginal property, the secrecy capacities of the channels with $\bar{F}_{\tilde{H}_r,\tilde{H}_e}(r, e)$ and $\bar{F}_{H_r,H_e}(r, e)$ are the same. Afterward, we show that $P(r \leq \tilde{H}_r \leq r + \epsilon, e \leq \tilde{H}_e) = 0$ with $r + \epsilon \leq e$ for arbitrarily small $\epsilon > 0$, given $H_r \geq_{st} H_e$. That is, given $H_r \geq_{st} H_e$, we can always find an equivalent channel where the square of Bob's channel realization is always larger than that of Eve's channel realization. Thus, the equivalent channel is a degraded one that has the same capacity as the original. We prove this property

[†]The Venn diagram in the conference version [27] is not correct in that the set of cx should not be included in $\mathcal{S}_{\mathcal{D}^+}$

[‡]They are equivalent in the sense that they have the same secrecy capacity.

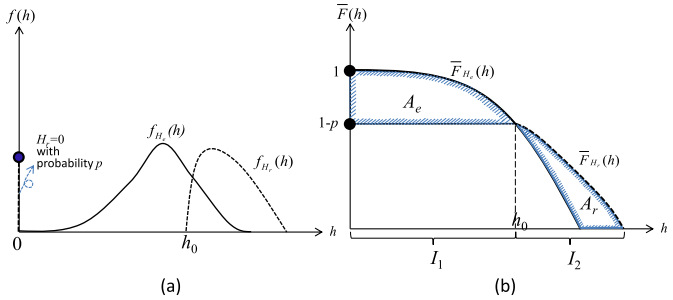

Fig. 2. An example where $H_r \not\geq_{st} H_e$ that results in a degraded wiretap channel. (a) PDFs of the two channels. Please note that $f_{H_r}(h)$ is a mixed distribution, which has probability $p > 0$ when $h = 0$. (b) CCDFs of the two channels.

as follows:

$$\begin{aligned} &P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e) \\ &\quad = P(r \leq \tilde{H}_r, e \leq \tilde{H}_e) - P(r+\epsilon \leq \tilde{H}_r, e \leq \tilde{H}_e) \\ &\quad \overset{(a)}{=} \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r,e) - \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r+\epsilon,e) \\ &\quad \overset{(b)}{=} \bar{F}_{\tilde{H}_e}(e) - \bar{F}_{\tilde{H}_e}(e) = 0, \end{aligned} \tag{8}$$

where (a) follows from the definition of the joint CCDF and (b) follows from (7) with the given property $H_r \geq_{st} H_e$, which implies that $\bar{F}_{\tilde{H}_r}(r) \geq \bar{F}_{\tilde{H}_r}(r+\epsilon) \geq \bar{F}_{\tilde{H}_e}(r+\epsilon) \geq \bar{F}_{\tilde{H}_e}(e)$. To ensure that $\tilde{h}_r \geq \tilde{h}_e$ for all random samples, we let $\epsilon \to 0$. Thus, as long as $H_r \geq_{st} H_e$, we can form an equivalent joint distribution that has the same marginal distribution as the original channel, such that the capacity is unchanged.

Now, we present a counterexample to prove that $H_r \geq_{st} H_e$ is not necessary to guarantee the existence of an equivalent degraded wiretap channel. It is clear that the example shown in Fig. 2 does not follow the usual stochastic order by definition. From the law of total probability, we can decompose the left-hand side of (8) as follows:

$$\begin{aligned} &P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e) \\ &= \underbrace{P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e | \tilde{H}_r \in I_1)}_{case1} P(\tilde{H}_r \in I_1) \\ &\quad + \underbrace{P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e | \tilde{H}_r \in I_2)}_{case2} P(\tilde{H}_r \in I_2), \end{aligned} \tag{9}$$

where $I_1 \triangleq \{h | 0 \leq h \leq h_0\}$ and $I_2 \triangleq \{h | h_0 < h\}$ and where h_0 is the other crossing point of $\bar{F}_{H_r}(h)$ and $\bar{F}_{H_e}(h)$ that is not $h = 0$. Because case 2 is identical to the first part of the proof, we directly find that $P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e | \tilde{H}_r \in I_2) = 0$. Now, we focus on case 1. By observing (7) and (8), we find that the right-hand side (RHS) of step (b) in (8) is not the only solution for obtaining $P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq H_e | \tilde{H}_r \in I_1) = 0$. That is, if $\bar{F}_{\tilde{H}_r,\tilde{H}_e}(r,e) = \bar{F}_{H_r}(r)$, $\bar{F}_{\tilde{H}_r,\tilde{H}_e}(r+\epsilon,e) = \bar{F}_{H_r}(r+\epsilon)$, and $\bar{F}_{H_r}(r) = \bar{F}_{H_r}(r+\epsilon)$, it is still possible to obtain $P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e) = 0$, i.e.,

$$P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq H_e) = \bar{F}_{H_r}(r) - \bar{F}_{H_r}(r+\epsilon) = 0. \tag{10}$$

For the first equality in (10) to hold, it is necessary that $\bar{F}_{H_r}(r) \leq \bar{F}_{H_e}(e)$ and $\bar{F}_{H_r}(r+\epsilon) \leq \bar{F}_{H_e}(e)$ when $e \in I_1$. In the example shown in Fig. 2, we can observe that the above two conditions are satisfied within the interval I_1. Furthermore, for the second equality in (10) to hold, $\bar{F}_{H_r}(r)$ must be constant within the interval I_1, which is also satisfied by the considered example. When $r + \epsilon = h_0$ occurs, because $e > r + \epsilon = h_0$, i.e., $e \in I_2$, and because of the fact that $\bar{F}_{H_r}(h) > \bar{F}_{H_e}(h)$ for all $h \in I_2$, it is clear that $P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e | \tilde{H}_r \in I_1) = 0$ from derivations identical to those in (8). As a result, we know that it is possible that a degraded wiretap channel with the same capacity as the original channel may still exist when $H_r \not\geq_{st} H_e$. This completes the proof. □

Remark 1: When the usual stochastic order is strict for any non-zero interval, it implies $\mathcal{S}_{\mathcal{D}^+}$. From Theorem 2, it can be easily proven that the intersection of $\mathcal{S}_{st}$ and $\mathcal{S}_{cx}$ occurs when $\bar{F}_{H_r}(h) = \bar{F}_{H_e}(h)$ for all h, which results in $C_s = 0$. Please note that the usual stochastic order is not a subset of $\mathcal{S}_{\mathcal{D}^+}$ because the usual stochastic order can result in a zero secrecy capacity.

Remark 2: Because the requirements for degradedness are stricter than less noisy and more capable [24], when $H_r \geq_{st} H_e$ is valid, by Lemma 1, the wiretap channel is also less noisy and more capable.

Note that Lemma 1 does not reveal the relation between the degradedness and the other stochastic orders. In the following lemmas, we characterize the relation first by means of an impossibility result in Lemma 2 and then by means of partial characterizations in Lemma 3 and Lemma 4.

Lemma 2: The increasing convex order is not sufficient to guarantee $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$, i.e., $\mathcal{S}_{\mathcal{D}^+} \cap \mathcal{S}_{icx} \neq \emptyset$ and $\mathcal{S}_{\mathcal{D}^+} \not\supseteq \mathcal{S}_{icx}$. $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$ does not necessarily imply $H_r \geq_{icx} H_e$, i.e., $\mathcal{S}_{icx} \not\supseteq \mathcal{S}_{\mathcal{D}^+}$.

The proof is given in Appendix II.

Lemma 3: If $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$, then $H_r \not\geq_{cx} H_e$. If $H_r \geq_{cx} H_e$, then $(H_r, H_e) \notin \mathcal{S}_{\mathcal{D}^+}$. Thus, $\mathcal{S}_{\mathcal{D}^+} \cap \mathcal{S}_{cx} = \emptyset$.

The proof is given in Appendix III.

In the following, we present a scenario in which $H_r \geq_{icx} H_e$ is a necessary condition for $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$.

Lemma 4: Assume that there is only one crossing point between $\bar{F}_{H_r}$ and $\bar{F}_{H_e}$ at $h = h_0$ and that $\bar{F}_{H_r}(h) > \bar{F}_{H_e}(h)$ for all $h > h_0$ and $\bar{F}_{H_r}(h) \leq \bar{F}_{H_e}(h)$ for all $h \leq h_0$. If H_r and H_e form a degraded wiretap channel and $H_r \not\geq_{icx} H_e$, then the ergodic secrecy capacity is zero.

Proof: First, we can construct the case $H_r \not\geq_{icx} H_e$ by setting $\int_0^{h_0}(-\bar{F}_{H_r}(h)+\bar{F}_{H_e}(h))dh > \int_{h_0}^{h_1}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))dh$, where h_0 is the crossing point between $\bar{F}_{H_r}(h)$ and $\bar{F}_{H_e}(h)$; recall that $h_1 = \max\{\text{supp}\,(\bar{F}_{H_r}(h)), \text{supp}(\bar{F}_{H_e}(h))\}$. Then, similar to the proof of the convex order in Lemma 2, we find that $C_s = 0$. □

Remark 3: Because of the availability of the full CSI at the receiver, the scenario considered in Fig. 2 can be treated as a wiretap channel with Bob's channel being in the On-Off mode, which can be used to model the analog front-end sensitivity at Bob. More specifically, Bob can only sense a signal with a magnitude larger than a certain threshold. Otherwise, Bob will treat the received signal as zero.

Remark 4: If we exchange the labels of H_r and H_e in Fig. 2 under the assumption that $A_e > A_r$, where A_e and A_r are defined in the proof of Lemma 2 then $H_r \geq_{icv} H_e$ is satisfied, i.e., $\int_{-\infty}^{h} \bar{F}_{H_r}(u)du \geq \int_{-\infty}^{h} \bar{F}_{H_e}(u)$ for all h. Note

that in this case, we may not be able to claim the existence of degradedness by the formulation of the joint CCDF in (7)*, and we cannot use the proof below* (9)*. However, by substituting* $U = X$ *into* (3)*, we can easily derive that if* $H_r \geq_{icv} H_e$*, then the ergodic secrecy rate is non-negative. This property can be extended to the more general case of* $H_r \geq_{Lt} H_e$ *[28]: if* $H_r \geq_{Lt} H_e$*, then* $E[\log(1+H_r)] \geq E[\log(1+H_e)]$*, which can be further interpreted as* $R_s \geq 0$ *by Gaussian signaling. Because* $C_s \geq R_s$*, we can claim that if* $H_r \geq_{Lt} H_e$*, then* $C_s \geq 0$.

IV. The Fast Fading Wiretap Channel With Statistical CSIT

Under conditions of statistical CSIT, for cases in which we are not able to identify equivalent degraded wiretap channels, Gaussian codebooks may not be optimal [11]. Thus, in this section, we instead propose to generate sub-channels via layered signaling. The underlying concept of layered signaling is to mimic the operation of the case with perfect CSIT such that Alice can choose sub-channels for transmission depending on the ergodic secrecy rate of each sub-channel even if Alice has only statistical CSI. At the beginning of this section, we verify the performance of the fast fading wiretap channel with statistical CSIT where the channel gain is binary expanded as multiple layers. We prove that independent coding over each sub-channel can yield the secrecy capacity in this scenario. We then apply this concept to fast fading Gaussian wiretap channels with statistical CSIT where $(H_r, H_e) \notin \mathcal{S}_{\mathcal{D}^+}$. We present several examples of practical fast fading channels with positive ergodic secrecy capacities. We also derive a general upper bound for this case. Finally we propose a layered signaling.

A. The Layered Erasure Wiretap Channel

In this section, we adapt the scheme from [21] to analyze the performance of the fast fading wiretap channel given the deterministic model shown in Fig. 3. Based on the deterministic model [22], we assume that Alice transmits a vector of q bits that are i.i.d. Bernoulli random sequences. Each bit is one layer of the transmitted signal. The fading effect is modeled by the erasure of the less significant bits, which cannot be transmitted successfully. The numbers of bits that can be received by Bob and Eve without being erased are $N_r[t]$ and $N_e[t]$, respectively, from the most significant bit (MSB) of the transmit signal. $N_r[t]$ and $N_e[t]$ correspond to PMFs of $P_{N_r}(n)$ and $P_{N_e}(n)$, respectively, and $t = 1, 2, \cdots$ is the time index. We also assume that Bob and Eve have perfect knowledge of their channels. Under the i.i.d. and memoryless assumptions for the channel states, the time index can be neglected to simplify the expression.

Theorem 3: The secrecy capacity of a layered erasure wiretap channel with statistical CSIT and full CSIR is

$$C_s = \sum_{n \in I_r} \bar{F}_{N_r}(n) - \bar{F}_{N_e}(n), \tag{11}$$

where $I_r = \{n | \bar{F}_{N_r}(n) > \bar{F}_{N_e}(n)\}$*,* $\bar{F}_{N_r}(n)$ *is the CCDF of Bob's channel, and* $\bar{F}_{N_e}(n)$ *is the CCDF of Eve's channel.*

The proof is given in Appendix IV.

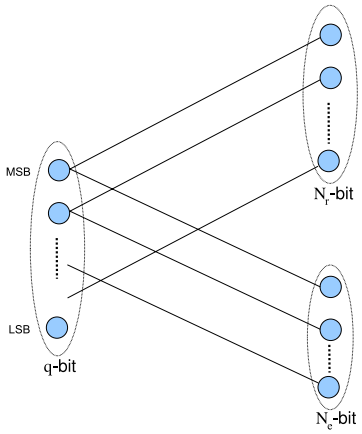

Fig. 3. The system model for the layered erasure wiretap channel.

B. The Fast Fading Gaussian Wiretap Channel With Statistical CSIT

Based on the stochastic orders introduced in Section II, in this sub-section, we first discuss several practical examples of the set $\mathcal{S}_{\mathcal{D}^+}$. From the capacity result for the layered signaling scheme presented in Section IV-A, we then introduce two capacity upper bounds for the considered channel. Based on our observations regarding the capacity upper bound, we propose an achievable scheme and derive the achievable ergodic secrecy rate for cases in which the wiretap channels do not belong to $\mathcal{S}_{\mathcal{D}^+}$.

1) Examples of the Set $\mathcal{S}_{\mathcal{D}^+}$*:* Based on Lemma 1, several examples with practical fading channels are provided in the following to explain how to determine the existence of a positive ergodic secrecy capacity given the distributions of the fading channels.

Example 1: Assume that the magnitudes of Bob's and Eve's channels are two independent Rayleigh random variables with variances of σ_r^2 and σ_e^2, respectively. Assume that $\sigma_r^2 > \sigma_e^2$. Because the square of the Rayleigh-distributed random variable X with variance σ_X^2 is exponentially distributed with the CCDF as $\exp(\frac{-x}{2\sigma_X^2})$, $\bar{F}_{H_r}(x) > \bar{F}_{H_e}(x)$ for all x. Then, from Lemma 1, we know that the capacity is given in (4). Note that this result coincides with [12, Lemma 1] for a single antenna.

Example 2: Assume that the magnitudes of Bob's and Eve's channels are two independent Nakagami-m random variables with shape parameters of m_r and m_e, respectively, and spread parameters of w_r and w_e. Then, the squares of their amplitudes follow an Erlang distribution. From Lemma 1, we know that the capacity is given in (4) if

$$\gamma\left(m_r, \frac{m_r}{w_r}x\right)\Gamma(m_e) \geq \gamma\left(m_e, \frac{m_e}{w_e}x\right)\Gamma(m_r) \text{ for all } x > 0, \tag{12}$$

where $\gamma(.)$ is the incomplete gamma function and $\Gamma(.)$ is the ordinary gamma function, defined as $\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}dt$. An example for which the inequality in (12) is valid is $(m_r, w_r) = (3, 2)$ with $(m_e, w_e) = (1, 2)$.

Example 3: Assume that the magnitudes of Bob's and Eve's channels are Nakagami-m and Rayleigh random variables, respectively. From Lemma 1, we know that the capacity is

given in (4) if

$$\gamma\left(m_r, \frac{m_r}{w_r}x\right)/\Gamma(m_r) \geq \exp\left(\frac{-x}{2\sigma_e^2}\right) \text{ for all } x > 0.$$

For example, $(m_r, w_r) = (3, 2)$ and $\sigma_e^2 = 3$ satisfy the requirement and result in the channels belonging to $\mathcal{S}_{\mathcal{D}^+}$.

2) Upper Bounds on the Ergodic Secrecy Capacity: In the following, we provide an explicit upper bound on the capacity of the fast fading Gaussian wiretap channel with statistical CSIT. If the channel distribution has non-zero probability at zero magnitude, we can further derive a tighter upper bound.

Theorem 4: The capacity of a fast fading Gaussian wiretap channel with power constraint $E[X^2] \leq 1$ and statistical CSIT is bounded from above by

$$C_s \leq C_s^{UB} = \frac{1}{2}\left\{\int_{h_r \in I_r} \log(1+h_r) f_{H_r}(h_r) dh_r - \int_{h_e \in I_r} \log(1+h_e) f_{H_e}(h_e) dh_e\right\}, \quad (13)$$

where $I_r = \{h | \bar{F}_{H_r}(h) > \bar{F}_{H_e}(h)\}$.

Proof: After channel enhancement such that $\bar{F}_{\tilde{H}_r}(h) = \max\{\bar{F}_{H_r}(h), \bar{F}_{H_e}(h)\}$ for all h, Bob's new channel is better than the original one. In addition, from Lemma 1, we have a degraded wiretap channel. Thus, we can use $I(X; Y_r|\tilde{H}_r) - I(X; Y_e|H_e)$ as the capacity upper bound on the original channel. Based on the proof of Theorem 1, we know that Gaussian input is optimal. Then, we can write the capacity upper bound as

$$C_s^{UB} = \frac{1}{2} E_{\tilde{H}_r, H_e}[\log(1+\tilde{H}_r) - \log(1+H_e)], \quad (14)$$

which can be further represented as in (13). □

A tighter upper bound can be found, as described in the following example, in the case of one crossing point if H_r has non-zero probability at $H_r = 0$. More specifically, under these conditions, we can apply a better enhancement scheme to obtain an equivalent degraded wiretap channel.

Lemma 5: For $H_r \geq_{icx} H_e$ with one crossing point between $\bar{F}_{H_r}$ and $\bar{F}_{H_e}$ and a non-zero probability $p > 0$ at $H_r = 0$, a tighter upper bound on the ergodic secrecy capacity can be obtained by applying the following enhancement:

$$\bar{F}_{\tilde{H}_r}(h) = \begin{cases} 1-p, h < \bar{F}_{H_e}^{-1}(1-p), \\ \max\{\bar{F}_{H_r}(h), \bar{F}_{H_e}(h)\}, h \geq \bar{F}_{H_e}^{-1}(1-p), \end{cases} \quad (15)$$

where $\bar{F}_{H_e}^{-1}$ is the inverse function of $\bar{F}_{H_e}$.

Proof: The fundamental idea behind the proof is to perform the necessary enhancement but not excessive enhancement, as in Theorem 4. If there is a non-zero probability p for $H_r = 0$, then we can follow the same steps as in the derivation of (9) to construct the degraded channel. More specifically, within the interval where $\bar{F}_{H_e}(h)$ is larger than $\bar{F}_{H_r}(h)$, i.e., $h < h_0$, instead of using the enhancement $\bar{F}_{\tilde{H}_r}(h) = \max\{\bar{F}_{H_r}(h), \bar{F}_{H_e}(h)\}$, it is instead sufficient to use $\bar{F}_{\tilde{H}_r}(h) = 1-p$ for all $h < \bar{F}_{H_e}^{-1}(1-p)$ to construct the degraded channel. For $h \geq \bar{F}_{H_e}^{-1}(1-p)$, we continue to use $\bar{F}_{\tilde{H}_r}(h) = \max\{\bar{F}_{H_r}(h), \bar{F}_{H_e}(h)\}$. Because $1-p \leq \max\{\bar{F}_{H_r}(h), \bar{F}_{H_e}(h)\}$ for all $h \in (0, \bar{F}_{H_e}^{-1}(1-p))$, it is easy to see that this enhancement results in a tighter C_s^{UB} than that of Theorem 4. □

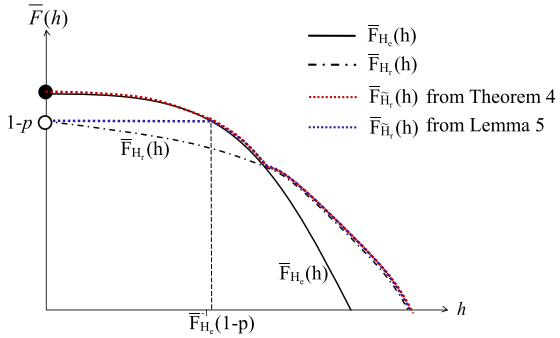

Fig. 4. Comparison of the upper bounds from Theorem 4 and Lemma 5.

An illustrative example is shown in Fig. 4. From this example, it is readily apparent that the upper bound from Lemma 5 is tighter than that from Theorem 4.

3) The Achievable Ergodic Secrecy Rate: From our analysis of the upper bound, we observe that if we do not transmit within the interval of h where Eve's channel has a higher CCDF than does Bob's, we will have a higher ergodic secrecy rate. Using a Gaussian codebook, we are unable to achieve this goal because there is no concept of channel partitioning. To overcome the shortcoming of using Gaussian signaling, we consider a binary expansion scheme [22] for the uniform transmit signal for the following reason. To improve the performance, we attempt to construct sub-channels such that we can avoid the transmission of messages on sub-channels with zero secrecy rates. The concept of *layers* in the binary expansion scheme is one possible means of achieving our goal.

After applying binary expansion [22] to the transmit signal, we obtain

$$Y = \sqrt{H}X + Z = \sum_{n=1}^{\infty} \sqrt{H}\left(\sqrt{3}X_n 2^{-n}\right) + Z, \quad (16)$$

where H denotes the square of the channel magnitude, Z is AWGN with zero mean and unit variance, and $X_n = 2b_n - 1$ is the n-th digit of X, with $b_n \sim Bern(1/2)$. We assume that $b_1, b_2, \cdots$ are i.i.d. It is evident that X is continuous and uniformly distributed on $\{-1, 1\}$. The factor of $\sqrt{3}$ is introduced to ensure that the transmit signal satisfies the unit channel input power constraint introduced in Sec. II because the variance of $\sum_{n=1}^{\infty} X_n 2^{-n}$ is $1/3$. By the assumption of full CSI at the receiver, we can rewrite (16) without changing the capacity as follows:

$$Y^{(h)} = \sum_{n=1}^{\infty} X_n 2^{-n} + \frac{Z}{\sqrt{3H}}, \quad (17)$$

$$= V_n + U_n + \frac{Z}{\sqrt{3H}} \quad (18)$$

where

$$V_n \triangleq \sum_{j=1}^{n} X_j 2^{-j}, \quad (19)$$

TABLE I
PARAMETERS AND NUMERICAL RESULTS FOR LAYERED SIGNALING

(m_r, w_r)	$E_{H_r}[h(X+Z_r^{(h)})]$	$E_{H_r}[h(V_1'^{(h)})]$	$E_{H_r}[h(V_2'^{(h)})]$	$E_{H_r}[h(V_3'^{(h)})]$
(10,1)	1.764	1.4317	1.6905	1.7451
(5,1)	1.7746	1.4485	1.7036	1.7575
(1,1)	1.8953	1.6034	1.8307	1.8796
(m_e, w_e)	$E_{H_e}[h(X+Z_e^{(h)})]$	$E_{H_e}[h(V_1'^{(h)})]$	$E_{H_e}[h(V_2'^{(h)})]$	$E_{H_e}[h(V_3'^{(h)})]$
(1,1.2)	1.8581	1.5496	1.79	1.8418
(1,1.4)	1.8278	1.5047	1.7563	1.8108
(0.2,1)	3.8633	3.6738	3.8201	3.853
(1,1.6)	1.8023	1.4664	1.7278	1.7846

$$U_n \triangleq \sum_{j=n+1}^{\infty} X_j 2^{-j}. \tag{20}$$

Here, $U_n \sim Unif(-2^{-n}, 2^{-n})$ is independent of V_n and Z. Bob's and Eve's channels are therefore modeled as $Y_r^{(h)} = X + Z_r/\sqrt{3H_r} \triangleq X + Z_r^{(h)}$ and $Y_e^{(h)} = X + Z_e/\sqrt{3H_e} \triangleq X + Z_e^{(h)}$, respectively. In the following, we present our result for the ergodic secrecy rate based on the layered transmission scheme.

Theorem 5: The following ergodic secrecy rate is achievable:

$$R_s = \sum_{n\in\{k|R_k>0\}} R_n, \tag{21}$$

where the ergodic secrecy rate of the n-th layer is

$$R_n = \left(E_{H_r}\left[h\left(X+Z_r^{(h)}\right) - h\left(V_n'^{(h)}\right)\right] - E_{H_e}\left[h\left(X+Z_e^{(h)}\right) - h\left(V_n'^{(h)}\right)\right]\right)^+ \tag{22}$$

and $V_n'^{(h)}$ follows the PDF

$$f_{V_n'^{(h)}}(v) = \sum_{k=0}^{2^{n-1}-1} \left\{\bar{F}_G^{(h)}\left(v - 1 + 2^{-(n-2)}k + 2^{-n}\right) - \bar{F}_G^{(h)}\left(v - 1 + 2^{-(n-2)}k + 3\cdot 2^{-n}\right)\right\}, \tag{23}$$

with $\bar{F}_G^{(h)}$ being the CCDF of the Gaussian random variable with zero mean and variance $1/(3h)$.

Proof: We first calculate the n-th layer of the first term on the RHS of (3) with $U = X$ as

$$\begin{aligned} & I\left(X_n; Y_r^{(h)}\middle| H_r\right) \\ &= h\left(Y_r^{(h)}\middle| H_r\right) - h\left(Y_r^{(h)}\middle| X_n, H_r\right) \\ &\overset{(a)}{=} h\left(X+Z_r^{(h)}\middle| H_r\right) - h\left(V_{n-1}+U_n+Z_r^{(h)}\middle| H_r\right) \\ &\overset{(b)}{\triangleq} h\left(X+Z_r^{(h)}\middle| H_r\right) - h\left(V_{n-1}+V_{r,n}^{(h)}\middle| H_r\right) \\ &\overset{(c)}{\triangleq} h\left(X+Z_r^{(h)}\middle| H_r\right) - h\left(V_{r,n}'^{(h)}\middle| H_r\right) \\ &= E_{H_r}\left[h\left(X+Z_r^{(h)}\middle| H_r=h\right) - h\left(V_{r,n}'^{(h)}\middle| H_r=h\right)\right], \end{aligned} \tag{24}$$

where in (a), we use the definitions (17), (18), (19), and (20), and in (b), we define $V_{r,n}^{(h)} \triangleq U_n + Z_r^{(h)}$, which has the following PDF:

$$f_{V_{r,n}^{(h)}}(v) = 2^{n-1}\left(\bar{F}_G^{(h)}\left(v - \frac{1}{2^n}\right) - \bar{F}_G^{(h)}\left(v + \frac{1}{2^n}\right)\right), \tag{25}$$

derived in Appendix V. Note that the PDF of $X + Z_r^{(h)}$ can be straightforwardly derived by substituting $n = 0$ into (25), i.e., $V_{r,0}^{(h)} = U_0 + Z_r^{(h)} = X + Z_r^{(h)}$ with the PDF

$$f_{Y_r^{(h)}} = 2^{-1}\left(\bar{F}_G^{(h)}(v-1) - \bar{F}_G^{(h)}(v+1)\right). \tag{26}$$

In the notation of $V_{r,n}^{(h)}$, the superscript (h) represents that the quantity is a function of the given h and the subscript n denotes that it is of the n-th layer. In (c), we define $V_{r,n}'^{(h)} \triangleq V_{n-1} + V_{r,n}^{(h)}$, which has the PDF

$$f_{V_{r,n}'^{(h)}}(v) = \frac{1}{2^{n-1}} \sum_{k=0}^{2^{n-1}-1} f_{V_{r,n}^{(h)}}(v - 1 + 2^{-(n-2)}k + 2^{-(n-1)}), \tag{27}$$

as shown in Appendix VI.

Similarly, we can represent the n-th layer from Alice to Eve as

$$\begin{aligned} & I\left(X_n; Y_e^{(h)}\middle| H_e\right) \\ &= E_{H_e}\left[h\left(X+Z_e^{(h)}\middle| H_e=h\right) - h\left(V_{e,n}'^{(h)}\middle| H_e=h\right)\right]. \end{aligned} \tag{28}$$

After subtracting (28) from (24) and exploiting the fact that the PDFs of $V_{r,n}'^{(h)}$ and $V_{e,n}'^{(h)}$ are the same for a given h and are denoted by $f_{V_n'}^{(h)}$, we can further rearrange the ergodic secrecy rate expression as in (22). Finally, after substituting (25) into (27), we obtain (23), which completes the proof. □

Remark 5: We observe that the rate expression in (22) is similar to the fast fading ergodic

secrecy capacity with full CSIT [29], i.e., $C_s = \frac{1}{2}E_{H_r,H_e}\left[(\log(1+P(h_r,h_e)H_r)-\log(1+P(h_r,h_e)H_e))^+\right]$, *where* $P(h_r,h_e)$ *is the power allocation based on the knowledge of* h_r *and* h_e *at the transmitter, in the sense that in both cases, the summation/expectation is over positive terms. More specifically, in [29], each sub-channel corresponds to one specific channel gain after quantization, which is similar to the distinction between each layer in the layered signaling scheme. That is, for these two scenarios, Alice can determine the sub-channel on which to transmit based on either statistical CSIT or perfect CSIT.*

Remark 6: Note that from (21)*, it seems that the layered signaling scheme predominates over normal Gaussian signaling because layer selection can be performed in the former scheme. However, this is not precisely true because in the layered signaling scheme, the channel input distribution is assumed to be uniform for the ease of binary decomposition, which may not be optimal. It can be seen from the numerical results that sometimes Gaussian signaling may outperform the proposed layered signaling scheme.*

Remark 7: Discrete signaling, e.g., the quadrature amplitude modulation (QAM) scheme used in [11], is a special case of the proposed layered scheme in the complex field. More specifically, the transmission of M-pulse amplitude modulation (PAM) signaling with $M = 2^n, n \in \mathbb{N}$, *is equivalent to the proposed scheme with only the first n layers being transmitted while layers with indices larger than n are discarded. For example, if we set* $n = 2$ *in* (18) *and discard* U_2*, we obtain* $X = \{-3/4, -1/4, 1/4, 3/4\}$ *with equal probability, which is identical to a 4-PAM signaling scheme. Moreover, QAM constellations can be straightforwardly obtained by extending the domain of* X *in* (18) *from real to complex.*

V. Numerical Results

In this section, we consider several examples to demonstrate the performance of the proposed upper and lower bounds. As a baseline for comparison, we use $X \sim N(0,1)$, with an ergodic secrecy rate of $R_s^G = \frac{1}{2}(E_{H_r}[\log(1+H_r)] - E_{H_e}[\log(1+H_e)])^+$. Note that obtaining an analytical form by further bounding (22) may not be sufficiently tight. That is, to guarantee weak secrecy, we need to bound the first and second expectations in (22) from below and above, respectively, which may be insufficient to ensure positive secrecy rates. Thus, we resort to numerically computing the differential entropies in (22) based on the known PDFs of X, V_{n-1}, $V_{r,n}^{(h)}$, $V_{e,n}^{(h)}$, $Z_r^{(h)}$, and $Z_e^{(h)}$. The calculated results for the considered scenario are listed in Table I. For simplicity, up to 3 layers are considered. Therefore, the numerically computed achievable rate is a lower bound on the proposed ergodic secrecy rate. In the first example, we assume that both Bob's and Eve's channels are Nakagami-m distributed with parameters of $m_r = 10$ and $w_r = 1$ and of $m_e = 1$ and $w_e = 1.2$, respectively. From Fig. 5, it is clear that this wiretap channel does not follow the usual stochastic order because of the crossing point. However, we can use Theorem 4 to obtain an upper bound on the ergodic

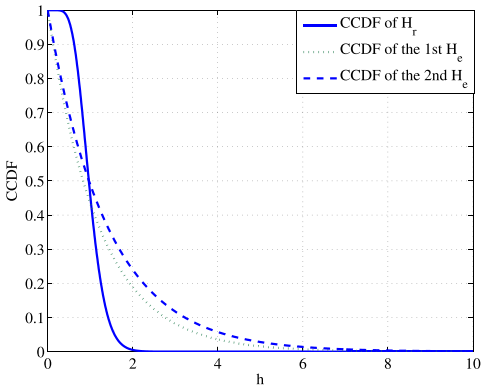

Fig. 5. The CCDF considered for the wiretap channel with $m_r = 10$, $w_r = 1$, $(m_{e_1}, w_{e_1}) = (1, 1.2)$ and $(m_{e_2}, w_{e_2}) = (1, 1.4)$.

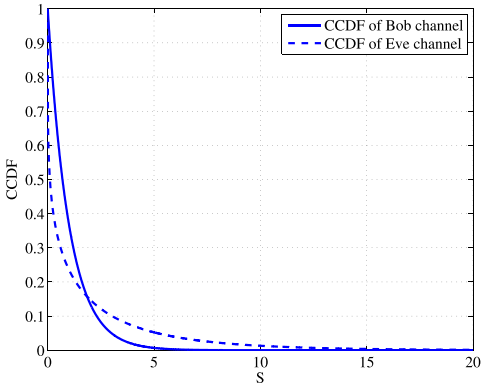

Fig. 6. The CCDF considered for the wiretap channel with $m_r = 1$, $w_r = 1$, $m_e = 0.2$ and $w_e = 1$.

secrecy capacity by transmitting only in the intervals in which Bob's channel has a larger CCDF than that of Eve's channel; this bound is 0.095 bits/channel use (bpcu). Using numerical integration, the proposed scheme is found to be bounded from below at 0.0318 bpcu. By contrast, Gaussian signaling results in a secrecy rate of $R_s^G = 0.0057$ bpcu, which is much lower than that of the proposed achievable scheme. In the same figure, we show the case in which $m_r = 10$, $w_r = 1$, $m_e = 1$ and $w_e = 1.4$. Note that in this case, $R_s^G = 0$, and our proposed upper and lower bounds are 0.0786 and 0.0131 bpcu, respectively. Because the region of a positive secrecy rate is enlarged, the connectivity of the network can be significantly improved by the proposed scheme. In another example, we set $m_r = 1$, $w_r = 1$, $m_e = 0.2$ and $w_e = 1$, with the CCDF as shown in Fig. 6. The proposed upper and lower bounds are 0.1625 and 0.1292 bpcu, respectively, where $R_s^G = 0.1247$. In this case, although the achievable ergodic secrecy rate of the proposed scheme is only slightly larger than that of Gaussian signaling, it is still closer to the upper bound.

To further demonstrate the advantages of using layered signaling, we present the following example. For channels with $m_r = 5$, $w_r = 1$, $m_e = 1$ and $w_e = 1.4$, we observe that the second-layer wiretap channel contributes negative values to the ergodic secrecy rate. Thus, by discarding the second layer,

we can obtain a higher achievable rate. Another example is $m_r = 10$ $w_r = 1$, $m_e = 1$ and $w_e = 1.6$, in which case only the third layer provides a positive ergodic secrecy rate. Therefore, using layered signaling, we can partially achieve the same effect as in the case with full CSIT although we have only statistical CSIT here.

VI. Conclusion

In this paper, we partially characterized the relation between the stochastic orders among Bob's and Eve's channels and the existence of degradedness and a positive ergodic secrecy capacity for a fast fading wiretap channel with only statistical channel state information for both channels at Alice. For more general orders, we also derived upper bounds and a capacity lower bound achieved through layered signaling. Finally, numerical results were presented to illustrate that in the case of Nakagami-m channels, the proposed achievable scheme outperforms the Gaussian codebook in several cases. More specifically, because the region of a positive secrecy rate is enlarged, the connectivity of the network can be significantly improved. Extensions to the multi-antenna and multi-user cases are worth investigating in our future studies.

Appendix I
Proof of Theorem 1

Proof: With only statistical CSIT, we cannot perform power allocation in the time domain. Thus, for each code symbol/channel realization, the power is the same. Therefore, we can write the secrecy capacity of the degraded wiretap channel as follows:

$$\max_{p_X:E[X^2]\le 1} I(X;Y_r|H_r) - I(X;Y_e|H_e) = \max_{p_X:E[X^2]\le 1} E[I(X;Y_r|H_r=h_r) - I(X;Y_e|H_e=h_e)].$$

For each channel realization h_r and h_e, we know from [7] that Gaussian input is optimal. After substituting the Gaussian input X into $I(X;Y_r|H_r) - I(X;Y_e|H_e)$, we obtain (4). □

Appendix II
Proof of Lemma 2

Proof: Note that there are only two requirements on the CCDFs in the example presented in Fig. 2: first, there is one crossing point between $\bar{F}_{H_r}(h)$ and $\bar{F}_{H_e}(h)$, and second, $\bar{F}_{H_r}(h)$ is constant when h is smaller than the crossing point h_0. Under the condition $A_e \le A_r$, where $A_e \triangleq \int_0^{h_0} |\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h)|dh$ and $A_r \triangleq \int_{h_0}^{\infty} |\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h)|dh$, this example satisfies $H_r \ge_{icx} H_e$. When we restrict the condition to $A_e = A_r$, we obtain $H_r \ge_{cx} H_e$. In the following, we use Fig. 2 to illustrate a counterexample.

We first prove the existence of $(H_r, H_e) \in \mathcal{S}_{\mathcal{D}^+}$ such that the increasing convex order is satisfied. Because the degradedness of Fig. 2 was already proven in Lemma 1, here we simply prove the existence of (H_r, H_e) such that $C_s > 0$. Then, we can first rewrite the ergodic secrecy capacity in (4) as follows:

$$\begin{aligned} C_s &= \frac{1}{2}\left\{E_{H_r}[\log(1+H_r)] - E_{H_e}[\log(1+H_e)]\right\} \\ &\overset{(a)}{=} \int \log(1+h) f_{H_r}(h)dh - \int \log(1+h) f_{H_e}(h)dh \\ &\overset{(b)}{=} \frac{1}{2}\log(1+h)\left(F_{H_r}(h) - F_{H_e}(h)\right)\Big|_0^{\infty} \\ &\quad + \frac{\log e}{2}\int_0^{\infty}\left(-F_{H_r}(h) + F_{H_e}(h)\right)\frac{1}{1+h}dh \\ &\overset{(c)}{=} \frac{1}{2}\log(1+h)\left(-\bar{F}_{H_r}(h) + \bar{F}_{H_e}(h)\right)\Big|_0^{\infty} \\ &\quad + \frac{\log e}{2}\int_0^{\infty}\left(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h)\right)\frac{1}{1+h}dh \\ &\overset{(d)}{=} \frac{\log e}{2}\int_0^{\infty}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h}dh \\ &\overset{(e)}{=} \frac{\log e}{2}\left\{\int_0^{h_0}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h}dh \right. \\ &\qquad \left. + \int_{h_0}^{\infty}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\} \\ &\triangleq C_L + C_R, \end{aligned} \tag{29}$$

where (a) follows from the definition of the expectation, (b) follows from integration by parts, and (c) follows from $\bar{F} = 1 - F$. (d) arises from the fact that the first term of the RHS of (c) is zero, which can be derived as follows. Note that $\bar{F}(h)\cdot\log(1+h)|_{h\to\infty} = P(H > h)\cdot\log(1+h)|_{h\to\infty}$ by the definition of the CCDF. Because $H \ge 0$, we can apply Markov's inequality to obtain

$$P(H > h)\cdot\log(1+h) \le \frac{E[H]}{h}\cdot\log(1+h).$$

Then, by applying $h \to \infty$ to the RHS of the above with the aid of L'Hôpital's rule, we find that $\lim_{h\to\infty}\frac{E[H]}{h}\cdot\log(1+h) = 0$ if $E[H]$ is finite. Note that a finite $E[H]$ is a meaningful and rather loose condition for a non-pathological channel in practice. In (e), we decompose the integration into two parts by assuming that there is only one crossing point h_0. Then, we can bound the two terms from below as follows:

$$\begin{aligned} C_R &\overset{(a)}{\ge} \frac{\log e}{2}\int_{h_0}^{h_1}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h_1}dh \\ &= \frac{\log e}{2}\frac{1}{1+h_1}\int_{h_0}^{h_1}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))dh = \frac{\log e}{2}\frac{1}{1+h_1}A_r, \end{aligned} \tag{30}$$

$$C_L \overset{(b)}{\ge} \frac{\log e}{2}\int_0^{h_0}(\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))dh = \frac{-\log e}{2}A_e, \tag{31}$$

where in (a), because this is a proof of existence, we can choose the example with bounded H_r and H_e for ease of illustration. Note that $h_1 \triangleq \max\{\text{supp}\,(\bar{F}_{H_r}(h)), \text{supp}(\bar{F}_{H_e}(h))\}$ and that the inequality arises from the fact that when $h_0 \le h \le h_1$, $1/(1+h_1) \le 1/(1+h)$ and $\bar{F}_{H_r}(h) > \bar{F}_{H_e}(h)$. In (b), we use the fact that when $0 \le h \le h_0$, $\bar{F}_{H_r}(h) < \bar{F}_{H_e}(h)$ and $1/(1+h) \le 1$. Thus, from (30) and (31), we know that $A_r > (1+h_1)A_e$ is a sufficient condition to guarantee $C_s > 0$.

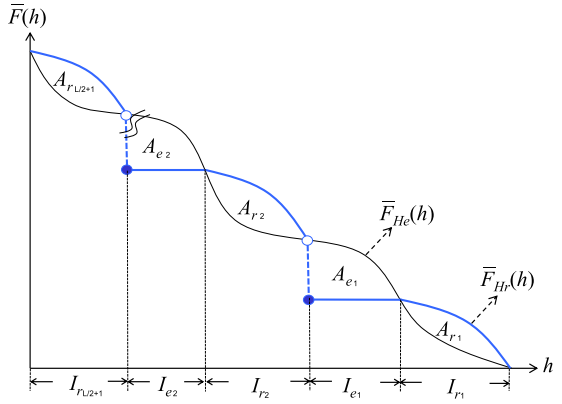

Fig. 7. An example where $H_r \not\geq_{icx} H_e$ but the channels are still degraded with a positive secrecy capacity.

The insight of this proof is that in the integration of (29), the term $\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h)$ is scaled by $1/(1+h)$. Thus, the "scaled" areas C_L and C_R may have the relation $|C_L| > |C_R|$ even if $A_e \leq A_r$, as long as h_0 is sufficiently large that A_r scales down sufficiently rapidly. We present a nontrivial example ($A_r = A_e$ is a trivial example), i.e., $A_r > A_e$, in the following. Assume that $\bar{F}_r(h) - \bar{F}_e(h) = -(\bar{F}_r(h+h_0) - \bar{F}_e(h+h_0))$ for $h \in [0, h_0]$; then, we have C_S as shown in (32), as shown at the bottom of this page, where (a) follows from the assumption that $\bar{F}_r(h) - \bar{F}_e(h) = -(\bar{F}_r(h+h_0) - \bar{F}_e(h+h_0))$. Note that $\bar{F}_{H_r}(h) < \bar{F}_{H_e}(h)$ when $h \in [0, h_0]$ and that $\bar{F}_{H_r}(h) > \bar{F}_{H_e}(h)$ when $h > 2h_0$, which is from $A_r > A_e$. Then, we can see that if h_0 is sufficiently large, then $1/(1+h+h_0)$ will be sufficiently small compared with $1/(1+h)$. Meanwhile, if $A_r - A_e$ is not too large, then the value of the second integration, which is positive, will be smaller than the absolute value of the first integration, which is negative. Then, we can obtain $C_S = 0$.

Now we prove the second part of this lemma, i.e., that the wiretap channel may still be degraded with a positive secrecy capacity even if the assumption of the increasing convex order is violated. A general example is illustrated in Fig. 7. The proof is similar to that of Lemma 1, with the following modifications. Assume that there are L crossing points, where L is even. Define the areas between the CCDFs $\bar{F}_{H_r}$ and $\bar{F}_{H_e}$ within the L intervals I_{r_1}, I_{e_1}, I_{r_2}, I_{e_2}, ..., $I_{r_{L/2+1}}$ as A_{r_1}, A_{e_1}, A_{r_2}, A_{e_2}, ..., $A_{r_{L/2+1}}$, respectively. Now, we impose the following two constraints on A_{r_k} and A_{e_k}:

c1. $A_{e_1} > A_{r_1}$.

c2. $A_{r_{L/2+1}}$ is sufficiently large to ensure $C_s > 0$.

From **c1**, we know that $H_r \geq_{icx} H_e$ is not valid by the definition of the increasing convex order. The validity of **c2** can be easily observed from the expression for C_s on the RHS of (b) in (29), which is rewritten here as

$$C_s = \frac{\log e}{2} \int_0^\infty (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h)) \frac{1}{1+h} dh. \tag{33}$$

Thus, to complete the proof of the second part, we need only show that the considered wiretap channel results in a degraded wiretap channel. Recall the method of constructing an equivalent wiretap channel in (7). We wish to prove that

$$\mathcal{P} \triangleq P(r \leq \tilde{H}_r \leq r+\epsilon, e \leq \tilde{H}_e) = 0, \tag{34}$$

given $r + \epsilon \leq e$ with arbitrarily small $\epsilon > 0$. Note that by the law of total probability, we can split (34) into $(L+1)^2$ cases, corresponding to all possible combinations of $(\tilde{H}_r, H_e)$ in these $L+1$ intervals. However, we need only check the following four cases: 1. $(\tilde{H}_r \in I_{r_j}, \tilde{H}_e \in I_{r_k})$, where $k \leq j$; 2. $(\tilde{H}_r \in I_{e_j}, \tilde{H}_e \in I_{e_k})$, where $k \leq j$; 3. $(\tilde{H}_r \in I_{r_j}, \tilde{H}_e \in I_{e_k})$, where $k < j$; and 4. $(\tilde{H}_r \in I_{e_j}, \tilde{H}_e \in I_{r_k})$, where $k \leq j$. In all other cases, $\mathcal{P} = 0$ because $\tilde{H}_r > \tilde{H}_e$. For case 1, $\tilde{H}_r \geq_{st} \tilde{H}_e$. Thus, from Lemma 1, we know that $\mathcal{P} = 0$. For case 2, $\mathcal{P} = \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r,e) - \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r+\epsilon, e) = \bar{F}_{\tilde{H}_e}(e) - \bar{F}_{\tilde{H}_e}(e) = 0$ if $k < j$ and $\mathcal{P} = \bar{F}_{\tilde{H}_r}(r) - \bar{F}_{\tilde{H}_r}(r+\epsilon) = 0$ if $k = j$. For case 3, $\mathcal{P} = \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r,e) - \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r+\epsilon, e) = \bar{F}_{\tilde{H}_e}(e) - \bar{F}_{\tilde{H}_e}(e) = 0$. For case 4, $\mathcal{P} = \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r,e) - \bar{F}_{\tilde{H}_r,\tilde{H}_e}(r+\epsilon, e) = \bar{F}_{\tilde{H}_e}(e) - \bar{F}_{\tilde{H}_e}(e) = 0$.

Thus, from the above, we can summarize that even if the icx assumption is violated, it is still possible to have a degraded wiretap channel with a positive secrecy capacity. □

APPENDIX III
PROOF OF LEMMA 3

Proof: Without loss of generality, assume that there are $2L-1$ crossing points between the two CCDFs for $h \in (0, \max\{\text{supp}\,(\bar{F}_{H_r}(h)), \text{supp}(\bar{F}_{H_e}(h))\})$. The reason we do not consider even numbers of crossing points is that in that case, the convex order cannot be satisfied by definition. Because we

$$\begin{aligned}
C_S &= \frac{\log e}{2}\left\{\int_0^{h_0}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh + \int_{h_0}^{\infty}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\} \\
&= \frac{\log e}{2}\left\{\int_0^{h_0}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh + \int_{h_0}^{2h_0}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh + \int_{2h_0}^{\infty}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\} \\
&= \frac{\log e}{2}\left\{\int_0^{h_0}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh + \int_0^{h_0}(\bar{F}_{H_r}(h+h_0)-\bar{F}_{H_e}(h+h_0))\frac{1}{1+h+h_0}dh \right. \\
&\qquad \left. + \int_{2h_0}^{\infty}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\} \\
&\overset{(a)}{=} \frac{\log e}{2}\left\{\int_0^{h_0}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\left(\frac{1}{1+h}-\frac{1}{1+h+h_0}\right)dh + \int_{2h_0}^{\infty}(\bar{F}_{H_r}(h)-\bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\},
\end{aligned} \tag{32}$$

are considering the convex order, the rightmost area between the two CCDFs should be A_{r_1}, followed by A_{e_1}, then A_{r_2}, etc., alternating in this manner. An illustrative example is provided in Fig. 7. Let the crossing point between the two adjacent areas A_{r_k} and A_{e_k} be denoted by h_k for $k = 1, \ldots, 2L-1$. Note that $h_1 > h_2 > \ldots > h_{2L-1}$. Also let the intervals on h to which A_{e_k} and A_{r_k} belong be denoted by I_{e_k} and I_{r_k}, respectively. Following the notations used in Appendix II, we can re-express the conditions for the convex order in Theorem 2 as follows:

$$\sum_{k=1}^{T} A_{r_k} \geq \sum_{k=1}^{T} A_{e_k}, T = 1 \cdots L-1, \tag{35}$$

$$\sum_{k=1}^{L} A_{r_k} = \sum_{k=1}^{L} A_{e_k}, \tag{36}$$

which is similar to the definition of majorization [30]. Here, however, the sequences $\{A_{r_k}\}$ and $\{A_{e_k}\}$ are not in the convex order. Thus, we are not able to use tools from majorization theory such as the Schur convexity, and we require the following derivation. We refer to the condition in (35) as the Tth condition and refer to that in (36) as the Lth condition. We can express the first condition as

$$A_{r_1} - A_{e_1} = \Delta_1, \tag{37}$$

where $\Delta_1 > 0$. Then, by substituting it into (35) and (36), we obtain

$$\sum_{k=2}^{T} A_{r_k} + \Delta_1 \geq \sum_{k=2}^{T} A_{e_k}, \quad T = 2 \cdots L-1, \tag{38}$$

$$\sum_{k=2}^{L} A_{r_k} + \Delta_1 = \sum_{k=2}^{L} A_{e_k}. \tag{39}$$

Afterward, we substitute each of the $L-2$ conditions in (38) into (39) to obtain

$$\sum_{k=S}^{L} A_{r_k} \leq \sum_{k=S}^{L} A_{e_k}, \quad S = 3 \cdots L. \tag{40}$$

By defining $\Delta_k \triangleq A_{r_k} - A_{e_k}$, we can further rearrange (39) and (40) as follows:

$$\Delta_2 = -\Delta_1 + \sum_{k=3}^{L} (-\Delta_k), \tag{41}$$

$$\Delta_S \leq \sum_{k=S+1}^{L} (-\Delta_k), \quad S = 3 \cdots L. \tag{42}$$

Then, we can bound C_s as (43) at the bottom of this page:

Now, by substituting (37), (41), and (42) into (43), we obtain

$$\begin{aligned}
C_s &= \frac{\log e}{2}\left\{\frac{\Delta_2}{1+h_2} + \left(\sum_{k=3}^{L}\frac{\Delta_k}{1+h_k}\right) + \frac{\Delta_1}{1+h_1}\right\} \\
&\overset{(a)}{=} \frac{\log e}{2}\left\{\frac{-1}{1+h_2}\left(\sum_{k=3}^{L}\Delta_k + \Delta_1\right)\right. \\
&\qquad \left. + \left(\sum_{k=3}^{L}\frac{\Delta_k}{1+h_k}\right) + \frac{\Delta_1}{1+h_1}\right\} \\
&\overset{(b)}{=} \frac{\log e}{2}\left\{\left(\sum_{k=3}^{L} a_{k,2}\Delta_k\right) + a_{1,2}\Delta_1\right\} \\
&\overset{(c)}{=} \frac{\log e}{2}\left\{a_{3,2}\Delta_3 + \left(\sum_{k=4}^{L} a_{k,2}\Delta_k\right) + a_{1,2}\Delta_1\right\} \\
&\overset{(d)}{\leq} \frac{\log e}{2}\left\{a_{3,2}\sum_{k=4}^{L}(-\Delta_k) + \left(\sum_{k=4}^{L} a_{k,2}\Delta_k\right) + a_{1,2}\Delta_1\right\} \\
&\overset{(e)}{=} \frac{\log e}{2}\left\{\left(\sum_{k=4}^{L} a_{k,3}\Delta_k\right) + a_{1,2}\Delta_1\right\} \\
&\overset{(f)}{=} \frac{\log e}{2}\left\{a_{L,L-1}\Delta_L + a_{1,2}\Delta_1\right\} \\
&\overset{(g)}{\leq} 0,
\end{aligned} \tag{44}$$

$$\begin{aligned}
C_s &= \frac{\log e}{2}\sum_{k=1}^{L}\left\{\int_{h\in I_{r_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h}dh + \int_{h\in I_{e_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h}dh\right\} \\
&\leq \frac{\log e}{2}\sum_{k=1}^{L}\left\{\int_{h\in I_{r_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h_k}dh + \int_{h\in I_{e_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))\frac{1}{1+h_k}dh\right\} \\
&= \frac{\log e}{2}\sum_{k=1}^{L}\frac{1}{1+h_k}\left\{\int_{h\in I_{r_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))dh + \int_{h\in I_{e_k}} (\bar{F}_{H_r}(h) - \bar{F}_{H_e}(h))dh\right\} \\
&= \frac{\log e}{2}\sum_{k=1}^{L}\frac{1}{1+h_k}\left\{A_{r_k} - A_{e_k}\right\} \\
&= \frac{\log e}{2}\sum_{k=1}^{L}\frac{\Delta_k}{1+h_k}.
\end{aligned} \tag{43}$$

where (a) follows from (41). In (b), we define $a_{k,j} \triangleq \frac{1}{1+h_k} - \frac{1}{1+h_j}$, $k, j = 1, \ldots L$. Note that $a_{k,j} > 0$ when $k > j$ because $h_k < h_j$. In (c), we extract the Δ_k with the smallest subscript k from the summation. In (d), we use (42). In (e), we again use the definition of $a_{k,j}$, i.e., $a_{k,i} = a_{k,j} - a_{i,j}$. In (f), we iteratively apply steps (c), (d), and (e) up to $k = L$. In (g), we use the fact that $a_{L,L-1} > 0$, $\Delta_L \leq 0$ according to (42), $a_{1,2} < 0$, and $\Delta_1 > 0$ by definition in (37), which completes the proof. □

Appendix IV
Proof of Theorem 3

Proof: This proof is adapted from [21]. For completeness, we explicitly restate the important steps here. The achievable scheme is to transmit the n-th bit X_n, which is i.i.d. Bernoulli with $p = 1/2$ in layer n. In the achievable scheme, we select $U_n = X_n$ and substitute it into (3) for each n. Then, the capacity from Alice to Bob on layer n, $n = 1 \cdots q$, is

$$H\left(\frac{1}{2}\right) \cdot P(N_r \geq n) + 0 \cdot P(N_r < n) = P(N_r \geq n) = \bar{F}_{N_r}(n).$$

Similarly, the capacity from Alice to Eve on each layer n is $\bar{F}_{N_e}(n)$. Therefore, the achievable secrecy rate of the wiretap channel on layer n is $\bar{F}_{N_r}(n) - \bar{F}_{N_e}(n)$. With independent coding on each level, the expected number of bits transmitted without erasure can be expressed as the RHS of (11), which is the summed rate of all layers belonging to the set I_r, i.e., only layers with positive rates. To prove the upper bound, we first enhance Bob's channel as follows:

$$\bar{F}_{\tilde{N}_r}(n) = \max\{\bar{F}_{N_r}(n), \bar{F}_{N_e}(n)\}, \quad \text{for all} \quad n. \tag{45}$$

Clearly, this enhancement scheme results in $\tilde{N}_r \geq_{st} N_e$. Then, by Lemma 1, we know that we have a degraded wiretap channel, and from [2], we know that the discrete memoryless wiretap channel capacity is $\max_{p(x)} I(X; Y_r) - I(X; Y_e)$. Then, we can derive the capacity upper bound by substituting $X = X^q$, $Y_r = X^{\tilde{N}_r}$, and $Y_e = X^{N_e}$ into the above capacity as follows:

$$\begin{aligned}
C_s^{UB} &= \max_{p(x^q)} I(X^q; X^{\tilde{N}_r}) - I(X^q; X^{N_e}) \\
&\overset{(a)}{=} \max_{p(x^q)} I(X^q; \tilde{N}_r, X^{\tilde{N}_r}) - I(X^q; N_e, X^{N_e}) \\
&\overset{(b)}{=} \max_{p(x^q)} I(X^q; \tilde{N}_r) + I(X^q; X^{\tilde{N}_r} | \tilde{N}_r) \\
&\quad - I(X^q; N_e) - I(X^q; X^{N_e} | \tilde{N}_e) \\
&\overset{(c)}{=} \max_{p(x^q)} I(X^q; X^{\tilde{N}_r} | \tilde{N}_r) - I(X^q; X^{N_e} | N_e) \\
&\overset{(d)}{=} \max_{p(x^q)} h(X^{\tilde{N}_r} | \tilde{N}_r) - h(X^{N_e} | N_e) \\
&\overset{(e)}{=} \max_{p(x^q)} \sum_{j=1}^{q} (\bar{F}_{\tilde{N}_r}(j) - \bar{F}_{N_e}(j)) h(X_j | X^{j-1}) \\
&\overset{(f)}{=} \sum_{j \in I_r} \bar{F}_{\tilde{N}_r}(j) - \bar{F}_{N_e}(j),
\end{aligned} \tag{46}$$

where (a) follows from $X^q \to X^{\tilde{N}_r} \to \tilde{N}_r$ and $X^q \to X^{\tilde{N}_e} \to \tilde{N}_e$, where $\tilde{N}_r$ is a deterministic function of $X^{\tilde{N}_r}$ and thus does not change the mutual information; (b) follows from the chain rule of mutual information; (c) follows from the fact that $\tilde{N}_r$ is independent of X^q because there is no instantaneous CSIT, thus $H(X^q|\tilde{N}_r) = E[H(X^q|\tilde{N}_r = \tilde{n}_r)] = 0$, and N_e has the same property; (d) follows from the fact $h(X^{\tilde{N}_r}|\tilde{N}_r, X^q) = h(X^{N_e}|N_e, X^q) = 0$ because given the transmit signal X^q and the fading states $\tilde{N}_r$ and N_e, we have exact knowledge of the receive signals $X^{\tilde{N}_r}$ and X^{N_e}, respectively; (e) follows from [21, Lemma 1(b)]; and (f) is obtained by selecting layers $j \in I_r$ because $j \in I_r^c$ does not contribute a positive rate. Note that we select the $\{X_j\}$ to be i.i.d. Bernoulli random variables with probability $p = 1/2$. By comparing the lower and upper bounds on the capacity, we conclude the proof. □

Appendix V
Proof of (25)

Proof: We first represent the PDF of U_n as $f_{U_n}(u) = 2^{n-1}\left(\mathfrak{u}(u + 2^{-n}) - \mathfrak{u}(u - 2^{-n})\right)$, where $\mathfrak{u}(.)$ denotes the unit step function. Because $V_{r,n}^{(h)} = U_n + Z_r^{(h)}$, the PDF of $V_{r,n}^{(h)}$ is

$$\begin{aligned}
&f_{V_{r,n}^{(h)}}(v) \\
&= \int_{-\infty}^{\infty} f_{U_n}(u) f_G^{(h)}(v-u) du \\
&\overset{(a)}{=} -f_{U_n}(u) F_G^{(h)}(v-u)\big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f'_{U_n}(u) F_G^{(h)}(v-u) du \\
&\overset{(b)}{=} -f_{U_n}(u) F_G^{(h)}(v-u)\big|_{-\infty}^{\infty} \\
&\quad + \int_{-\infty}^{\infty} (2^{n-1}\left(\delta(u+2^{-n}) - \delta(u-2^{-n})\right) F_G^{(h)}(v-u) du \\
&\overset{(c)}{=} \int_{-\infty}^{\infty} (2^{n-1}\left(\delta(u+2^{-n}) - \delta(u-2^{-n})\right) F_G^{(h)}(v-u) du \\
&\overset{(d)}{=} 2^{n-1}(F_G^{(h)}(v+2^{-n}) - F_G^{(h)}(v-2^{-n})) \\
&\overset{(e)}{=} 2^{n-1}(\bar{F}_G^{(h)}(v-2^{-n}) - \bar{F}_G^{(h)}(v+2^{-n})),
\end{aligned}$$

where $f_G^{(h)}$ is the PDF of $Z_r^{(h)}$; $F_G^{(h)}$ is the CDF of $Z_r^{(h)}$; (a) is obtained through integration by parts; in (b), we use the property that $\mathfrak{u}'(u) = \delta(u)$, where $\delta(u)$ is the delta function; (c) follows from the fact that $f_{U_n}(\infty) = f_{U_n}(-\infty) = 0$ and $0 \leq F_G^{(h)}(v-u) \leq 1$; (d) follows from the nature of the delta function; and (e) follows from the definition of the CCDF, $\bar{F}_G^{(h)}(v - 2^{-n}) = 1 - F_G^{(h)}(v - 2^{-n})$. □

Appendix VI
Proof of (27)

Proof: Note that for the n-th layer, V_{n-1} is used to calculate $V_{r,n}^{\prime(h)}$. Then, from (19), $V_{n-1} = \sum_{j=1}^{n-1} X_j 2^{-j}$, where X_j is generated by $2b_j - 1$ with $b_j \sim Bern(1/2), \forall j$. Thus, V_{n-1} produces the values $-1 + 2^{-(n-2)}k + 2^{-(n-1)}, k = 0 \cdots 2^{n-1} - 1$, with uniform probability $2^{-(n-1)}$. Then, we can represent the PDF as

$$f_{V_{n-1}}(v) = \frac{1}{2^{n-1}} \sum_{k=0}^{2^{n-1}-1} \delta(v - 1 + 2^{-(n-2)}k + 2^{-(n-1)}). \tag{47}$$

After convolving (47) with $f_{V_{r,n}^{(h)}}(v)$, we obtain (27). □

ACKNOWLEDGMENTS

The authors would like to thank the reviewers and the Associate Editor for their constructive suggestions, which greatly improved the presentation of the paper.

REFERENCES

[1] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Foundations and Trends in Communications and Information Theory: Information Theoretic Security*. Norwell, MA, USA: Now Publishers, 2009.
[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.
[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
[4] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.
[5] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.
[6] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.
[7] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.
[8] P. K. 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.
[9] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.
[10] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.
[11] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2792–2799, Sep. 2010.
[12] S.-C. Lin and P.-H. Lin, "On secrecy capacity of fast fading multiple-input wiretap channels with statistical CSIT," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 2, pp. 414–419, Feb. 2013.
[13] M. R. Bloch and J. N. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1840–1849, Sep. 2013.
[14] Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz), "A broadcast approach for fading wiretap channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.
[15] S.-C. Lin and C.-L. Lin, "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3293–3306, Jun. 2014.
[16] P. Mukherjee and S. Ulukus, "Fading wiretap channel with no CSI anywhere," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 1347–1351.
[17] J. Li and A. P. Petropulu, "On ergodic secrecy rate for Gaussian MISO wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.
[18] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.
[19] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
[20] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2009, pp. 2602–2606.
[21] D. N. C. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3453–3471, Jun. 2012.
[22] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
[23] G. Caire and S. Shamai (Shitz), "On the capacity of some channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.
[24] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[25] M. Shaked and G. Shanthikumar, *Stochastic Orders*. New York, NY, USA: Springer-Verlag, 2007.
[26] B. Hajek. (2014). *Notes for ECE 534: An Exploration of Random Processes for Engineers*. [Online]. Available: http://www.ifp.illinois.edu/~hajek/Papers/randomprocJan14.pdf
[27] P.-H. Lin and E. Jorswieck, "On the fading Gaussian wiretap channel with statistical channel state information at transmitter," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2014, pp. 121–126.
[28] A. Rajan and C. Tepedelenlioglu, "Stochastic ordering of fading channels through the Shannon transform," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1619–1628, Apr. 2015.
[29] 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.
[30] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, 2nd ed. New York, NY, USA: Academic, 1979.

Pin-Hsun Lin (M'10) received the B.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2000 and 2010, respectively. He was a Visiting Student with the University of Maryland, College Park, in 2009. From 2011 to 2012, he was with the INTEL-NTU Laboratory as a Postdoctoral Associate. Since 2012, he had been a Researcher with the Smart Wireless Laboratory, National Institute of Information and Communications Technology, Yokosuka, Japan. He is now a Postdoctoral Associate with the Communications Theory Laboratory, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, Germany. His research interests include wireless communications and information theory. He received the Best Ph.D. Dissertation Award.

Eduard Jorswieck (S'01–M'03–SM'08) received the Diplom-Ingenieur (M.S.) and Doktor-Ingenieur (Ph.D.) degrees in electrical engineering and computer science from the Technische Universitat Berlin, Germany, in 2000 and 2004, respectively. He was with the Broadband Mobile Communication Networks Department, Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Berlin, from 2000 to 2008. From 2005 to 2008, he was a Lecturer with the Technische Universitat Berlin. From 2006 to 2008, he was with the Department of Signals, Sensors and Systems, Royal Institute of Technology, as a Postdoctoral Researcher and an Assistant Professor. Since 2008, he has been the Head of the Chair of Communications Theory and a Full Professor with the Technische Universität Dresden, Germany. He is currently a Principal Investigator with the Excellence Cluster Center for Advancing Electronics Dresden and a Founding Member of the 5G Laboratory, Germany.

He has authored over 75 journal papers, seven book chapters, some 200 conference papers, and three monographs on these research topics. His main research interests are in the area of signal processing for communications and networks, applied information theory, and communications theory. He was a corecipient of the IEEE Signal Processing Society Best Paper Award in 2006 and coauthored papers that won the Best Paper or Best Student Paper Awards at the IEEE WPMC 2002, Chinacom 2010, the IEEE CAMSAP 2011, the IEEE SPAWC 2012, and the IEEE WCSP 2012.

Dr. Jorswieck was a member of the IEEE SPCOM Technical Committee (2008–2013), and is a member of the IEEE SAM Technical Committee. Since 2011, he has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. Since 2008, continuing until 2011, he has served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and a Senior Associate Editor until 2013. Since 2013, he has served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.