Secrecy Outage Capacity of Fading Channels

Onur Gungor, Jian Tan, Can Emre Koksal, Hesham El-Gamal, Ness B. Shroff

***Abstract*—This paper considers point to point secure communication over flat fading channels under an outage constraint. More specifically, we extend the definition of outage capacity to account for the secrecy constraint and obtain sharp characterizations of the corresponding fundamental limits under two different assumptions on the transmitter channel state information (CSI). First, we find the outage secrecy capacity assuming that the transmitter has perfect knowledge of the legitimate and eavesdropper channel gains. In this scenario, the capacity achieving scheme relies on opportunistically exchanging private keys between the legitimate nodes. These keys are stored in a key buffer and later used to secure delay sensitive data using the Vernam's one time pad technique. We then extend our results to the more practical scenario where the transmitter is assumed to know only the legitimate channel gain. Here, our achievability arguments rely on privacy amplification techniques to generate secret key bits. In the two cases, we also characterize the optimal power control policies which, interestingly, turn out to be a judicious combination of channel inversion and the optimal ergodic strategy. Finally, we analyze the effect of key buffer overflow on the overall outage probability.**

I. Introduction

Secure communication is a topic that is becoming increasingly important thanks to the proliferation of wireless devices. Over the years, several secrecy protocols have been developed and incorporated in several wireless standards; e.g., the IEEE 802.11 specifications for Wi-Fi. However, as new schemes are being developed, methods to counter the specific techniques also appear. Breaking this cycle is critically dependent on the design of protocols that offer provable secrecy guarantees. The information theoretic secrecy paradigm adopted here, allows for a systematic approach for the design of low complexity and provable secrecy protocols that fully exploit the intrinsic properties of the wireless medium.

Most of the recent work on information theoretic secrecy is, arguably, inspired by Wyner's wiretap channel [1]. In this setup, a passive eavesdropper listens to the communication between two legitimate nodes over a separate communication channel. While attempting to decipher the message, no limit is imposed on the computational resources available to the eavesdropper. This assumption led to defining **perfect secrecy capacity** as the maximum achievable rate subject to zero mutual information rate between the transmitted message and the signal received by the eavesdropper. In the additive Gaussian noise scenario [2], the perfect secrecy capacity turned out to be the difference between the capacities of the legitimate and eavesdropper channels. Therefore, if the eavesdropper channel has a higher channel gain, information theoretic secure communication is not possible over the main channel. Recent works have shown how to exploit multipath fading to avoid this limitation [3]–[5]. The basic idea is to opportunistically exploit the instants when the main channel enjoys a higher gain than the eavesdropper channel to exchange secure messages. This opportunistic secrecy approach was shown to achieve non-zero **ergodic secrecy capacity** even when **on average** the eavesdropper channel has favorable conditions over the legitimate channel. Remarkably, this result still holds even when the channel state information of the eavesdropper channel is not available at the legitimate nodes [3].

The ergodic result in [3] applies only to delay tolerant traffic, e.g., file downloads. Early attempts at characterizing the delay limited secrecy capacity drew the negative conclusion that non-zero delay limited secrecy rates are not achievable, over almost all channel distributions, due to **secrecy outage** events corresponding to the instants when the eavesdropper channel gain is larger than the main one [6], [7]. Later, it was shown in [8] that, interestingly, a non-zero delay limited secrecy rate could be achieved by introducing **private key queues** at both the transmitter and the receiver. These queues are used to store private key bits that are shared **opportunistically** between the legitimate nodes when the main channel is more favorable than the one seen by the eavesdropper. These key bits are used later to secure the delay sensitive data using the Vernam one time pad approach [9]. Hence, secrecy outages are avoided by simply storing the secrecy generated previously, in the form of key bits, and using them whenever the channel conditions are more advantageous for the eavesdropper. However, in [8], the authors do not provide sharp capacity results or derive the corresponding optimal power control policies, which is the main objective of this paper. In particular,

- We consider delay limited communication in a block fading channel where the messages to be transmitted in a block has to be communicated securely within that particular block. We find compact expressions of the secrecy outage capacity for the scenario where (i) perfect knowledge about the main and eavesdropper channels are available *a-priori* at the transmitter, referred to as *full channel state information (CSI)*, and (ii) only the perfect knowledge main channel states are available at the transmitter, referred to as *main CSI*. We provide a graphical approach to evaluate the capacity.
- We develop a (secrecy outage) capacity-achieving scheme

O. Gungor, C. E. Koksal, H. El-Gamal and N. B. Shroff are with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, 43210 (e-mail: gungoro@ece.osu.edu, koksal@ece.osu.edu, helgamal@ece.osu.edu, shroff@ece.osu.edu).

J. Tan is with the IBM T. J. Watson Research Center, Hawthorne, NY, 10532 (email: tanji@us.ibm.com)

This work was published in part in the Proceedings of INFOCOM 2010, San Diego, CA, and in the Proceedings of Conference on Information Sciences and Systems, CISS 2012, Princeton, NJ.

This work is supported in part by QNRF under grant NPRP 5-559-2-227, and by NSF under grants CNS-1054738, CNS-0831919, CCF-0916664.

that utilizes privacy amplification to generate secret key bits from the transmitted signal, and store them in the form of secret key bits in the transmitter and legitimate receiver. These key bits are utilized to secure the delay sensitive data using Vernam's one time pad. This approach is proven to be optimal even when the eavesdropper CSI is not known at the legitimate nodes, since the statistical knowledge of eavesdropper channel enables us to generate key bits over many fading blocks.
- We evaluate the optimal power allocation in order to achieve the secrecy outage capacity and provide a novel power controller, which combines secure waterfilling and channel inversion policies.
- Past studies that make use of a key queue assume that the associated buffer has an infinite size. Here, we analyze the impact of a finite buffer and explicitly evaluate the amount of reduction in the achievable secret data rate if a finite key buffer is used.

The rest of this paper is organized as follows. We formally introduce our system model in Section II. In Section III, we obtain the capacity results for the full and main CSI scenarios. The optimal power control policies, for both cases, are derived in Section IV. The effect of key buffer overflow on the outage probability is investigated in Section V. We provide simulations to support our main results in Section VI. Finally, Section VII offers some concluding remarks. To enhance the flow of the paper, the proofs are collected in the Appendices.

II. System Model

We study a point-to-point wireless communication link, in which a transmitter wishes to send information to a legitimate receiver, in the presence of a passive eavesdropper. We divide time into discrete slots, where blocks are formed of N channel uses, and B blocks combine to form a super-block. Let the communication period consist of S super-blocks. We use the notation (s,b) to denote the b^{th} block in the s^{th} super-block. We adopt a block fading channel model, in which the channel is assumed to be constant over a block, and changes randomly from one block to the next. Within each block (s,b), the observed signals at the receiver and at the eavesdropper are:

$$\mathbf{Y}(s,b) = G_m(s,b)\mathbf{X}(s,b) + \mathbf{W}_m(s,b)$$
$$\mathbf{Z}(s,b) = G_e(s,b)\mathbf{X}(s,b) + \mathbf{W}_e(s,b),$$

respectively, where $\mathbf{X}(s,b) \in \mathbb{C}^N$ is the transmitted signal, $\mathbf{Y}(s,b) \in \mathbb{C}^N$ is the received signal by the legitimate receiver, $\mathbf{Z}(s,b) \in \mathbb{C}^N$ is the received signal by the eavesdropper, and $\{\mathbf{W}_m(s,b)\}_{s=1,b=1}^{S,B}$ and $\{\mathbf{W}_e(s,b)\}_{s=1,b=1}^{S,B}$ are two mutually independent i.i.d. vector processes that are also independent of other random variables. Each sample of $\mathbf{W}_m(s,b) \in \mathbb{C}^N$ and $\mathbf{W}_e(s,b) \in \mathbb{C}^N$ are independently drawn from circularly symmetric, unit variance normal distribution. We assume that the channel gains of the main channel $G_m(s,b)$ and the eavesdropper channel $G_e(s,b)$ are i.i.d. complex random variables. The power gains of the fading channels are denoted by $H_m(s,b) = |G_m(s,b)|^2$ and $H_e(s,b) = |G_e(s,b)|^2$. We sometimes use the vector notation $\mathbf{H}(\cdot) = [H_m(\cdot)\ H_e(\cdot)]$ for simplicity, use the notation $\mathbf{H}^{s,b} = \{\mathbf{H}\}_{s'=1,b'=1}^{s,b}$ to denote the set of channel gains $\mathbf{H}(s',b')$ observed until block (s,b), and use backslash as relative complement operator, e.g., $\mathbf{H}^{S,B}\backslash\mathbf{H}(s,b)$ denotes the set of gains of all blocks except (s,b). We use identical notation for other parameters as well, and denote the sample realization sequences with lowercase letters. We assume that the probability density function of instantaneous channel gains, denoted as $f(\mathbf{h})$, is well defined, and is known by all parties. Under both full CSI and main CSI cases, the eavesdropper has complete knowledge of both the main and the eavesdropper channels. Let $P(s,b)$ denote the power allocated at block (s,b). We consider a long term power constraint (or average power constraint) such that,

$$\frac{1}{SB}\sum_{s=1}^{S}\sum_{b=1}^{B} P(s,b) \le P_{\text{avg}} \tag{1}$$

for some $P_{\text{avg}} > 0$.

Let $\{W(s,b)\}_{s=1,b=1}^{S,B}$ denote the set of messages to be transmitted with a delay constraint. $W(s,b)$ becomes available to the transmitter at the beginning of block (s,b), and needs to be securely communicated and decoded at the legitimate receiver at the end of that particular block. We consider the problem of constructing $(2^{NR}, N)$ codes to communicate message packets $W(s,b) \in \{1, \cdots, 2^{NR}\}$ of equal size, which consists of:

1) A stochastic encoder that maps $(w(s,b),\ \mathbf{x}^{s,b-1})$ to $\mathbf{x}(s,b)$ based on the available CSI, where $\mathbf{x}^{s,b-1}$ summarizes the previously transmitted signals[1], and
2) A decoding function that maps $\mathbf{y}^{s,b}$ to $\hat{w}(s,b)$ at the legitimate receiver.

Note that we consider the current block $\mathbf{x}(s,b)$ to be a function of the past blocks $\mathbf{x}^{s,b-1}$ as well. This kind of generality allows us to store shared randomness to be exploited in the future to increase the achievable secrecy rate.

Define the error event with parameter δ at block (s,b) as

$$E(s,b,\delta) = \big\{\hat{W}(s,b) \neq W(s,b)\big\} \cup \left\{\frac{1}{N}\|\mathbf{X}(s,b)\|^2 > P(s,b) + \delta\right\} \tag{2}$$

which occurs either when the decoder makes an error, or when the power expended is greater than $P(s,b) + \delta$. Let $W^{S,B}\backslash W(s,b)$ denote the messages to be communicated in all the blocks except $W(s,b)$. The equivocation rate at the eavesdropper is defined as the entropy rate of the message at block (s,b), conditioned on the received signal by the eavesdropper during the transmission period, available eavesdropper CSI, and messages[2] to be communicated in all blocks except the message at block (s,b), which is equal to $\frac{1}{N}H(W(s,b)|\mathbf{Z}^{S,B}, \mathbf{h}^{S,B}, W^{S,B}\backslash W(s,b))$. The secrecy outage event at rate R with parameter δ at block (s,b) is defined as

$$\mathcal{O}_{\text{sec}}(s,b,R,\delta) = \mathcal{O}_{\text{eq}}(s,b,R,\delta) \cup \mathcal{O}_{\text{inf}}(s,b,R,\delta) \tag{3}$$

[1]An exception is for $b = 1$, in which case the previous signals are summarized by $\mathbf{x}^{s-1,B}$.

[2]Although the messages $\{W(s,b)\}_{s=1,b=1}^{S,B}$ are mutually independent, they may be dependent conditioned on eavesdroppers' received signal $\mathbf{Z}^{S,B}$, therefore equivocation expression includes conditioning on $W^{S,B}\backslash W(s,b)$.

where the equivocation outage occurs if the equivocation rate at block (s,b) is less than $R-\delta$,

$$\mathcal{O}_{\text{eq}}(s,b,R,\delta) = \left\{ \frac{1}{N} H\big(W(s,b)|\mathbf{Z}^{S,B}, W^{S,B}\backslash W(s,b), \mathbf{h}^{S,B}\big) < R-\delta \right\} \tag{4}$$

and information outage occurs if accumulated mutual information on the message $W(s,b)$ remains below its entropy, $R-\delta$:

$$\mathcal{O}_{\text{inf}}(s,b,R,\delta) = \left\{ \frac{1}{N} I\left(W(s,b); \mathbf{Y}^{s,b}\right) < R-\delta \right\}. \tag{5}$$

Defining $\bar{\mathcal{O}}_{\text{sec}}(\cdot)$ as the complement of the event $\mathcal{O}_{\text{sec}}(\cdot)$, we now characterize the notion of ϵ-achievable secrecy capacity.

Definition 1: Rate R is achievable securely with at most ϵ probability of secrecy outage if, for any fixed $\delta > 0$, there exists a sequence of codes of rate no less than R such that, for all large enough S, B and N, the conditions

$$\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)) < \delta \tag{6}$$

$$\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta)) < \epsilon + \delta \tag{7}$$

are satisfied for all (s,b), $s \neq 1$.

We call such R an ϵ-achievable secrecy rate. Note that the conditioning in (4) is based on the realization $\mathbf{h}^{S,B}$ of all the channel gains, and the probability expressions are over $\mathbf{H}^{s,b}$. Also note that the security constraints are not imposed on the first super-block.

Definition 2: The ϵ-achievable secrecy capacity is the supremum of all ϵ-achievable secrecy rates.

Remark 1: The notion of secrecy outage was previous defined and used in [6], [7]. However, those works did not consider the technique of storing shared randomness for future use, and in that case, secrecy outage depends only on the instantaneous channel states, and hence the achievable data rates were rather suboptimal. In our case, secrecy outage depends on previous channel states as well. We illustrate the suboptimality of the previous works in Example 1. Note that we do not impose a secrecy outage constraint on the first superblock ($s = 1$). We refer to the first superblock as an initialization phase used to generate initial common randomness between the legitimate nodes. This phase only needs to appear *once* in the communication lifetime of that link. In other words, when a session (which consists of S superblocks) between the associated nodes is over, they would have sufficient number of common key bits for the subsequent session, and would not need to initiate the initialization step again.

III. Capacity Results

In this section, we investigate ϵ-achievable secrecy capacity under two different cases; full CSI and main CSI at the transmitter. We show in capacity proofs that the outage capacity achieving power allocation functions lie in the space of stationary power allocation functions that are functions of instantaneous transmitter CSI. Hence for **full CSI**, we constrain ourselves to the set $\mathcal{P}_F : \{\mathbf{h}\} \rightarrow \mathbb{R}^+ \cup \{0\}$ of stationary power allocation policies that are functions of $\mathbf{h} = [h_m\ h_e]$, Similarly for **main CSI** we consider the set $\mathcal{P}_M$ of power allocation policies that are functions of h_m only. For a given power allocation function $P \in \mathcal{P}_F$, define

$$R_m(\mathbf{h}, P(\mathbf{h})) \triangleq \log(1 + P(\mathbf{h})h_m) \tag{8}$$

$$R_s(\mathbf{h}, P(\mathbf{h})) \triangleq [\log(1 + P(\mathbf{h})h_m) - \log(1 + P(\mathbf{h})h_e)]^+ \tag{9}$$

where $[\cdot]^+ = \max(\cdot, 0)$, and the logarithms are with respect to base 2. Note that, $R_m(\cdot)$ is the supremum of achievable main channel rates, without the secrecy constraint. Also, $R_s(\cdot)$ is the non-negative difference between main channel and eavesdropper channel's supremum achievable rates. Similarly, for main CSI, we consider $R_m(\mathbf{h}, P(h_m))$ and $R_s(\mathbf{h}, P(h_m))$ for $P \in \mathcal{P}_M$.

A. Full CSI

Theorem 1: Let the transmitter have full CSI. Then, for any ϵ, $0 \leq \epsilon < 1$, the ϵ-achievable secrecy capacity is equal to $C_F(\epsilon)$ bits per channel use, where

$$C_F(\epsilon) = \max_{P \in \mathcal{P}_F} \frac{\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]}{1-\epsilon} \tag{10}$$

subject to:

$$\mathbb{P}\left(R_m(\mathbf{H}, P(\mathbf{H})) < \frac{\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]}{1-\epsilon} \right) \leq \epsilon \tag{11}$$

$$\mathbb{E}[P(\mathbf{H})] \leq P_{\text{avg}} \tag{12}$$

A detailed proof of achievability and converse part is provided in Appendix A. Here, we briefly justify the result. For a given power allocation function $P \in \mathcal{P}_F$, $R_s(\mathbf{h}, P(\mathbf{h}))$ is the supremum of the secret key generation rates within a block that experiences channel gains $\mathbf{h}$ [2]. This implies that the expected achievable secrecy rate [3] is $\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]$ without the outage constraint. With the outage constraint, the fluctuations of $R_s(\mathbf{H}, P(\mathbf{H}))$ due to fading are unacceptable, since $R_s(\mathbf{H}, P(\mathbf{H}))$ can go below the desired rate when the channel conditions are unfavorable (e.g., when $H_m < H_e$, $R_s(\mathbf{H}, P(\mathbf{H})) = 0$). Hence, we utilize secret key buffers to smoothen out these fluctuations to provide secrecy rate of $\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]$ at each block. The generated secrecy is stored in secret key buffers of both the transmitter and receiver, and is utilized to secure message of same size using Vernam's one-time pad technique. Note that every single generated key bit is used *exactly* once, such that keys generated in s-th superblock are used in $s+1$'st superblock. Secrecy outage may still occur when either there is not enough key bits left at the key queue, or the main channel rate for the block remains below the desired rate. In this case, we do not attempt to transmit the message, hence no key bits are expended. Therefore, with ϵ probability of secrecy outage, a secrecy rate of $\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]/(1-\epsilon)$ could be achieved. The channel outage constraint (11) on the other hand is a necessary condition for the main channel to support the desired rate, avoid information outages (5), and satisfy the secrecy outage constraint in (7).

Example 1: Consider a four state system, where H_m and H_e takes values from the set $\{1, 10\}$ and the joint probabilities are as given in Table I. Let the average power constraint be $P_{\text{avg}} = 0.5$, and there is no power control, i.e., $P(\mathbf{h}) = P_{\text{avg}}$ $\forall \mathbf{h}$. The achievable instantaneous secrecy rate, and the main channel rate at each state are given in Tables II and III, respectively. According to the pessimistic result in [6,8], no non-zero rate can be achieved with a secrecy outage probability $\epsilon < 0.6$ in this case. However, according to Theorem 1, rate $R = \frac{\mathbb{E}[R_s(\mathbf{H}, P_{\text{avg}})]}{1-\epsilon} = \frac{0.8}{1-\epsilon}$ can be achieved with ϵ secrecy outage probability[3] for any $\epsilon \geq 0.2$. A sample path is provided for both schemes in Figure 1, and it is shown how our scheme avoids secrecy outage in the second block. Note that, for $\epsilon < 0.2$, the rate $R = \frac{\mathbb{E}[R_s(\mathbf{H}, P_{\text{avg}})]}{1-\epsilon}$ cannot be achieved due the limitation of instantaneous main channel rate, as shown in Table III. Instead, a secrecy rate of only $R = 0.58$ can be achieved. In Example 2, we show that, through a more clever control of the power expended, we can achieve much higher rates.

TABLE I
$\mathbb{P}(\mathbf{h})$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	0.1	0.1
10	0.4	0.4

TABLE II
$R_s(\mathbf{h}, P_{\text{AVG}})$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	0	0
10	2	0

TABLE III
$R_m(\mathbf{h}, P_{\text{AVG}})$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	0.58	0.58
10	2.58	2.58

B. Main CSI

Theorem 2: Let the transmitter have main CSI. Then, for any ϵ, $0 \leq \epsilon < 1$, the ϵ-achievable secrecy capacity is equal to $C_M(\epsilon)$ bits per channel use, where

$$C_M(\epsilon) = \max_{P \in \mathcal{P}_M} \frac{\mathbb{E}[R_s(\mathbf{H}, P(H_m))]}{1-\epsilon} \tag{13}$$

subject to:

$$\mathbb{P}\left(R_m(\mathbf{H}, P(H_m)) < \frac{\mathbb{E}[R_s(\mathbf{H}, P(H_m))]}{1-\epsilon} \right) \leq \epsilon \tag{14}$$

$$\mathbb{E}[P(H_m)] \leq P_{\text{avg}} \tag{15}$$

Although the problems (10)-(12) and (13)-(15) are of the same form, due to the absence of eavesdropper CSI, the maximization in this case is over power allocation functions $\mathcal{P}_M$ that depend on the main channel state only. Hence, $C_M(\epsilon) \leq C_F(\epsilon)$. A detailed proof of achievability and converse is provided in Appendix B. As in the full CSI case, our achievable scheme uses similar key buffers and Vernam's one time pad technique to secure the message. The main difference is the generation of secret key bits. Due to the lack of knowledge of $H_e(s, b)$ at the transmitter, secret key bits cannot be generated within a block. In [8], a sub-optimal slot division approach was utilized, in which part of each slot was used in generating keys, and the other part was used in transmitting the delay sensitive data. Instead, we generate keys over super-blocks using privacy amplification, carefully designed based on the sample distribution of $H_e(s, b)$. Roughly, over a superblock the receiver can reliably obtain $NB\mathbb{E}[R_m(\mathbf{H}, P(H_m))]$ bits of information, while the eavesdropper can obtain $NB\mathbb{E}[R_m(\mathbf{H}, P(H_m)) - R_s(\mathbf{H}, P(H_m))]$ bits of information. With privacy amplification, $NB\mathbb{E}[R_s(\mathbf{H}, P(H_m))]$ bits of secret key can be extracted.

Now, we show that power allocation policy has minimal impact on the performance in the high power regime.

Theorem 3: For any $\epsilon > 0$, the ϵ-achievable secrecy capacities with full CSI and main CSI converge to the same value

$$\lim_{P_{\text{avg}} \to \infty} C_F(\epsilon) = \lim_{P_{\text{avg}} \to \infty} C_M(\epsilon) = \frac{\mathbb{E}\left[\log\left(H_m/H_e\right)\right]^+}{(1-\epsilon)} \tag{16}$$

Proof: For $\mathbf{h} \equiv [h_m \ h_e]$ such that $h_m > h_e$, we can see from (9) that $\lim_{P(\mathbf{h}) \to \infty} R_s(\mathbf{h}, P(\mathbf{h})) = \log\left(\frac{h_m}{h_e}\right)$, and for $h_m \leq h_e$, $R_s(\mathbf{h}, P(\mathbf{h})) = 0$. Furthermore, for $h_m > 0$, we can see from (8) that $\lim_{P(\mathbf{h}) \to \infty} R_m(\mathbf{h}, P(\mathbf{h})) = \infty$. Let $P(\mathbf{h}) = P_{\text{avg}}$ $\forall \mathbf{h}$ (no power control), which does not require any CSI. Then, we get

$$\lim_{P_{\text{avg}} \to \infty} \mathbb{E}[R_s(\mathbf{H}, P_{\text{avg}})] = \mathbb{E}\left[\log\left(H_m/H_e\right)\right]^+ < \infty. \tag{17}$$

Combining (16) and (17), we get

$$\lim_{P_{\text{avg}} \to \infty} \mathbb{P}\left(R_m(\mathbf{H}, P_{\text{avg}}) < \frac{\mathbb{E}[R_s(\mathbf{H}, P_{\text{avg}})]}{1-\epsilon} \right) = \mathbb{P}(H_m = 0)$$

and $\mathbb{P}(H_m = 0) = 0$, since probability density function of $\mathbf{H}$ is well defined. Hence, channel outage constraints (11) and (14) are not active in the high power regime. From (10)-(12) and (13)-(15), we conclude that (16) holds. ■

Our simulation results also illustrate that the power allocation policy has minimal impact on the importance in the high power regime. On the other hand, when the average power is limited, the optimality of the power allocation function is of critical importance, which is the focus of the following section.

IV. Optimal Power Allocation Strategy

A. Full CSI

The optimal power control strategy, $P^* \in \mathcal{P}_F$ is the stationary strategy that solves the optimization problem (10)-(12). In this section, we will show that P^* is a time-sharing between the channel inversion power policy, and the secure waterfilling policy. We first introduce the channel inversion power policy P_{inv}, which is the *minimum* required power to maintain main channel rate of R.

$$P_{\text{inv}}(h_m, R) = \frac{2^R - 1}{h_m} \tag{18}$$

[3]Although Theorem 1 is stated for the case where random vector $\mathbf{H}$ is continuous, the result similarly applies to discrete $\mathbf{H}$ as well.

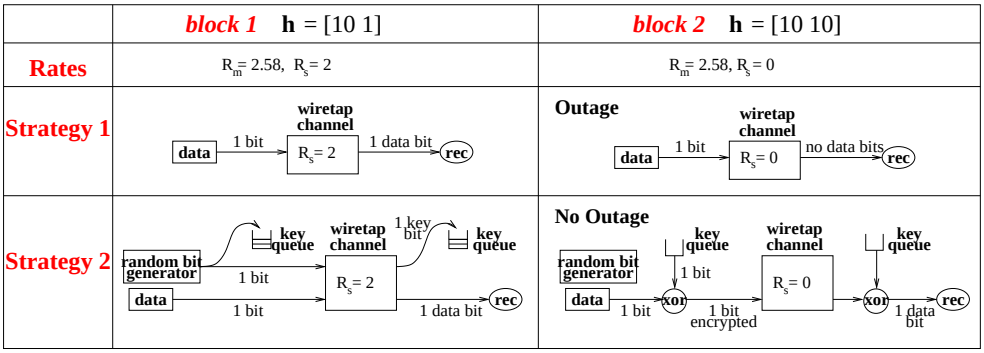

Fig. 1. A sample path. With strategy 2, secrecy outage can be avoided for block $t = 2$ via the use of key bits.

Note that, main CSI knowledge is sufficent for P_{inv}. Next we introduce P_{wf},

$$P_{\text{wf}}(\mathbf{h},\lambda) = \frac{1}{2}\Bigg[\sqrt{\left(\frac{1}{h_e}-\frac{1}{h_m}\right)^2+\frac{4}{\lambda}\left(\frac{1}{h_e}-\frac{1}{h_m}\right)} - \left(\frac{1}{h_e}+\frac{1}{h_m}\right)\Bigg]^+. \tag{19}$$

We call it the 'secure waterfilling' power policy because it maximizes the ergodic secrecy rate without any outage constraint, and resembles the 'waterfilling' power control policy. Here, the parameter λ determines the power expended on average. Now, let us define a time-sharing region

$$\mathcal{G}(\lambda,k) = \Big\{\mathbf{h} : [R_s(\mathbf{h},P_{\text{inv}}(h_m,R)) - R_s(\mathbf{h},P_{\text{wf}}(\mathbf{h},\lambda))]^+ - \lambda\,[P_{\text{inv}}(h_m,R) - P_{\text{wf}}(\mathbf{h},\lambda)]^+ \geq k\Big\} \tag{20}$$

which is a function of parameters λ and k.

Theorem 4: P^* is the unique solution to

$$P^*(\mathbf{h}) = P_{\text{wf}}(\mathbf{h},\lambda^*) + \mathbf{1}\big(\mathbf{h}\in\mathcal{G}(\lambda^*,k^*)\big)\big(P_{\text{inv}}(h_m,C_F(\epsilon)) - P_{\text{wf}}(\mathbf{h},\lambda^*)\big)^+ \tag{21}$$

subject to: $k^* \leq 0, \lambda^* > 0$

$$C_F(\epsilon) = \mathbb{E}[R_s(\mathbf{H},P^*(\mathbf{H}))]/(1-\epsilon) \tag{22}$$

$$\mathbb{P}(\mathbf{H}\in\mathcal{G}(\lambda^*,k^*)) = 1-\epsilon \tag{23}$$

$$\mathbb{E}[P^*(\mathbf{H})] = P_{\text{avg}} \tag{24}$$

Proof: Define a sub-problem

$$\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))] = \max_{P\in\mathcal{P}_F}\mathbb{E}\left[R_s(\mathbf{H},P(\mathbf{H}))\right] \tag{25}$$

subject to: $P(\mathbf{h}) \geq 0,\ \forall\mathbf{h}$

$$\mathbb{E}[P(\mathbf{H})] \leq P_{\text{avg}}, \tag{26}$$

$$\mathbb{P}\big(R_m(\mathbf{H},P(\mathbf{H})) < R\big) \leq \epsilon \tag{27}$$

Let P^R be the power allocation function that solves this sub-problem. Note that for $R = \mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]/(1-\epsilon)$, this problem is identical to (10)-(12), hence giving us $R = C_F(\epsilon)$, and $P^R = P^*$. We will prove the existence and uniqueness of such R.

Lemma 1: There exists a unique $R_{\max} > 0$ such that the sub-problem (25)-(27) has a solution for all $R \leq R_{\max}$, which is found by solving

$$P_{\text{avg}} = \int_{h_m\geq c} P_{\text{inv}}(h_m,R_{\max})f(\mathbf{h})d\mathbf{h} \tag{28}$$

for $\mathbf{h} \equiv [h_m\ h_e]$, where the constant c is chosen such that $\mathbb{P}(H_m \leq c) = \epsilon$.

Proof is provided in Appendix C-A.

Lemma 2: For any $R \leq R_{\max}$,

$$P^R(\mathbf{h}) = P_{\text{wf}}(\mathbf{h},\lambda) + \mathbf{1}\left(\mathbf{h}\in\mathcal{G}(\lambda,k)\right)\left(P_{\text{inv}}(h_m,R) - P_{\text{wf}}(\mathbf{h},\lambda)\right)^+ \tag{29}$$

where $k \in (-\infty,0]$ and $\lambda \in (0,+\infty)$ are parameters that satisfy (26) and (27) with equality.

Proof is provided in Appendix D. It is left to show there exists a unique R that satisfies $R = \mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]/(1-\epsilon)$.

Lemma 3: $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ is a continuous non-increasing function of R.

Proof is provided in Appendix C-B.

Lemma 4: There exists a unique R, $0 \leq R \leq R_{\max}$, which satisfies $R = \mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]/(1-\epsilon)$.

Proof is provided in Appendix C-C. This concludes the proof of the theorem. ■

Due to (21), the optimal power allocation function is a time-sharing between the channel inversion and secure waterfilling; a balance between avoiding channel outages, hence secrecy outages, and maximizing the expected secrecy rate. The time sharing region $\mathcal{G}(\lambda,k)$ determines the instants $\mathbf{h}$, for which avoiding channel outages are guaranteed through the choice of $P(\mathbf{h}) = \max(P_{\text{inv}}(h_m,R),P_{\text{wf}}(\mathbf{h},\lambda))$. (23) ensures that channel outage probability is at most ϵ, and (24) ensures that average power constraint is met with equality. (22), on the other hand, is an immediate consequence of (10).

Note that, an extreme case is $P^*(\mathbf{h}) = P_{\text{wf}}(\mathbf{h},\lambda^*)\ \forall\mathbf{h}$, which occurs when $P_{\text{inv}}(\mathbf{h},R) \leq P_{\text{wf}}(\mathbf{h},\lambda^*)$ for any $\mathbf{h} \in \mathcal{G}(\lambda^*,k^*)$, i.e., the secure waterfilling solution itself satisfies the channel outage probability in (11). However, that the other extreme ($P^*(\mathbf{h}) = P_{\text{inv}}(h_m,R^*)\ \forall\mathbf{h}$) cannot occur for any non-zero ϵ due to (21). The parameter $C_F(\epsilon)$ can be found graphically as shown in Figure 2, by plotting $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ and and $(1-\epsilon)R$ as a function of R. The abcissa of the unique intersection point is $R = C_F(\epsilon)$.

Example 2: Consider the same system model in Example 1. We have found that for $R = \frac{0.8}{1-\epsilon}$ bits/channel use is achievable

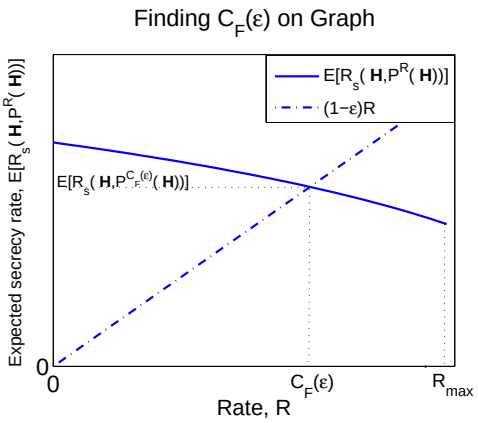

Fig. 2. Finding $C_F(\epsilon)$ with graphical approach

with ϵ probability of secrecy outage with no power control, i.e., $P(\mathbf{h}) = 0.5\ \forall \mathbf{h}$ for $\epsilon \geq 0.2$. Let $\epsilon = 0.2$, we will see if we can do better than $R = 1$ with power control. Solving the problem (21)-(24), we can see that[4] the time-sharing, and power expended in each state are as given in Tables IV and V. For $\mathbf{h} \equiv [h_m\ h_e] = [10\ 1]$, i.e., the legitimate channel has a better gain, secure waterfilling is used and when $\mathbf{h} = [10\ 10]$, secret key bits cannot be generated, but channel inversion is used to guarantee a main channel rate of R, which is secured by the excess keys generated during the state $\mathbf{h} = [10\ 1]$. As a result, we can see that a rate of $C_F(0.2) = 1.26$ bits per channel use is achievable, which corresponds to 26% increase with respect to no power control. As mentioned in Theorem 3, this gain diminishes at the high power regime, i.e., when $P_{\text{avg}} \to \infty$.

We also study the case with $\epsilon = 0$, for which a secrecy rate of $R = 0.58$ can be achieved, as illustrated in Example 1. Solving Problem (21)-(24) for $\epsilon = 0$, we obtain the power allocation in Table VII, for which a secrecy rate of $C_F(0) = 0.9$ bit per channel use is achievable. As shown in Table VI, channel inversion guarantees a main channel rate of $C_F(0)$ at all times, which was not possible without power control as shown in Example 1.

TABLE IV
TIME SHARING REGIONS, $\epsilon = 0.2$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	wf	wf
10	wf	inv

TABLE V
$P^*(\mathbf{h})$, $\epsilon = 0.2$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	0	0
10	1.11	0.14

TABLE VI
TIME SHARING REGIONS, $\epsilon = 0$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	inv	inv
10	wf	inv

TABLE VII
$P^*(\mathbf{h})$, $\epsilon = 0$

$\downarrow h_m \setminus h_e \rightarrow$	1	10
1	0.86	0.86
10	0.73	0.08

[4]Although Theorem 4 assumes $\mathbf{H}$ is a continuous random vector, the results similarly hold for the discrete case as well.

B. Main CSI

Here, we find the optimal power control strategy $P^* \in \mathcal{P}_M$, which solves the optimization problem (13)-(15). Let us define the main CSI secure waterfilling power policy P_w, such that $P_w(h_m, \lambda)$ is the maximum of 0, and the solution of the following equation

$$\frac{\partial \mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]}{\partial P(h_m)} = \frac{h_m \mathbb{P}(h_e \leq h_m)}{1 + h_m P(h_m)} - \int_0^{h_m} \left(\frac{h_e}{1 + h_e P(h_m)} \right) f(h_e) dh_e - \lambda = 0 \quad (30)$$

Theorem 5: $P^*(h_m)$ is the unique solution to

$$P^*(h_m) = P_w(h_m, \lambda^*) + \mathbf{1}(h_m \geq c) \left(P_{\text{inv}}(h_m, C_M(\epsilon)) - P_w(h_m, \lambda^*) \right)^+ \quad (31)$$

subject to: $\lambda^* > 0$

$$C_M(\epsilon) = \mathbb{E}[R_s(\mathbf{H}, P^*(H_m))]/(1-\epsilon) \quad (32)$$

$$\mathbb{P}(H_m \geq c) = 1 - \epsilon \quad (33)$$

$$\mathbb{E}[P^*(H_m)] = P_{\text{avg}} \quad (34)$$

where $\mathbb{E}[R_s(\mathbf{H}, P^*(H_m))]$ is the expected secrecy rate under the power allocation policy P^*.

Note that, optimal power allocation function takes a form similar to Theorem 4, except $P_w(h_m, \lambda)$ replaces $P_{\text{wf}}(\mathbf{h}, \lambda)$, and the time-sharing regions are different.

Proof: The proof follows the approach in Full CSI case, hence we omit the details for brevity. Define the sub-problem

$$\mathbb{E}[R_s(\mathbf{H}, P^R(H_m))] = \max_{P \in \mathcal{P}_M} \mathbb{E}\big[R_s(\mathbf{H}, P(H_m))\big] \quad (35)$$

subject to: $P(h_m) \geq 0,\ \forall h_m$

$$\mathbb{E}[P(H_m)] \leq P_{\text{avg}}, \quad (36)$$

$$\mathbb{P}\left(R_m(\mathbf{H}, P(H_m)) < R\right) \leq \epsilon \quad (37)$$

Let $P^R \in \mathcal{P}_M$ be the power allocation function that solves this sub-problem. Lemmas 1 and 4 also hold in this case. The only difference is the following lemma, which replaces Lemma 2 in Full CSI.

Lemma 5: For any $R \leq R_{\max}$ and h_m,

$$P^R(h_m) = P_w(h_m, \lambda) + \mathbf{1}(h_m > c) \left(P_{\text{inv}}(h_m, R) - P_w(h_m, \lambda) \right)^+$$

where c is a constant that satisfies $\mathbb{P}(H_m \geq c) = 1 - \epsilon$, and $\lambda \in (0, +\infty)$ is a constant that satisfies (36) with equality. The proof is similar to the proof of Lemma 2, and is provided in Appendix E. ■

The graphical solution in Figure 2 to find $C_F(\epsilon)$ also generalizes to the main CSI case.

V. SIZING THE KEY BUFFER

The proofs of the capacity results of Section III assume availability of *infinite size* secret key buffers at the transmitter and receiver, which mitigate the effect of fluctuations in the achievable secret key bit rate due to fading. Finite-sized buffers, on the other hand will lead to a higher secrecy outage probability due to wasted key bits by the key buffer overflows. Here, we revisit the full CSI problem, and consider key buffer

sizes *normalized* with respect to the number of channel uses in a block, N, as follows. We define $M(\epsilon,R)$ to be the normalized buffer size[5], in terms of bits per channel use, required to achieve rate R with at most ϵ probability of secrecy outage.

Theorem 6: Let $\epsilon'>\epsilon$, and $\kappa(x)\triangleq x\log(x)$. Then,

$$\lim_{\epsilon'\searrow\epsilon}\frac{M(\epsilon',C_F(\epsilon))}{\kappa\left(\frac{\mathbb{V}ar[R_s(\mathbf{H},P^{C_F(\epsilon)}(\mathbf{H}))]+(C_F(\epsilon))^2\epsilon(1-\epsilon)}{(\epsilon'-\epsilon)C_F(\epsilon)}\right)}\leq 1 \quad (38)$$

where $P^{C_F(\epsilon)}\in\mathcal{P}_F$ is the power allocation policy defined in (29), for parameter $R=C_F(\epsilon)$.

Before providing the proof, we first interpret this result. If buffer size is infinite, we can achieve rate $C_F(\epsilon)$ with ϵ probability of secrecy outage. With finite buffer, we can achieve the same rate with ϵ' probability of secrecy outage. Considering this difference to be the price that we have to pay due to the finiteness of the buffer, we can see that the normalized buffer size required scales with $\mathrm{O}\left(\frac{1}{\epsilon'-\epsilon}\log\frac{1}{\epsilon'-\epsilon}\right)$, as $\epsilon'-\epsilon\to 0$.

Proof: Achievability follows from simple modifications to the capacity achieving scheme described in Appendix A. We will first study the key queue dynamics, then using the heavy traffic limits, we provide an upper bound to the key loss ratio due to buffer overflows. Then, we relate key loss ratio to the secrecy outage probability, and conclude the proof.

For the key queue dynamics, we use a single index t to denote the time index instead of the double index (s,b), where $t=sB+b$. We consider transmission at outage secrecy rate of R, and use power allocation function P^R, which solves the problem (25)-(27). Let us define $\{Q_M(t)\}_{t=1}^{\infty}$ as the key queue process with buffer size M, and let $Q_M(1)=0$. To simplify notation, let us consider $R_s(t)\equiv R_s(\mathbf{h}(t),P^R(\mathbf{h}(t)))$ to denote the value of $R_s(\cdot)$ at block t, and similarly define $R_m(t)$ as well. Then, during each block t,

1) The transmitter and receiver agree on secret key bits at rate $R_s(t)$ bits /channel use using privacy amplification, and store the key on their secret key buffers.
2) The transmitter pulls key bits at rate R bits / channel use from its secret key buffer to secure the message stream at rate R bits/ channel use using one time pad, and transmits over the channel.

as explained in Appendix A. The last phase is skipped if outage ($\mathcal{O}_{\text{enc}}(t)$) is declared, which is triggered by one of the following events

- Channel Outage ($\mathcal{O}_{\text{ch}}(t)$): The channel cannot support reliable transmission at rate R, i.e. $R_m(t)<R$.
- Key Outage ($\mathcal{O}_{\text{key}}(t)$): There are not enough key bits in the key queue to secure the message at rate R. This event occurs when $Q_M(t)+R_s(t)-R<0$.
- Artificial outage ($\mathcal{O}_{\text{a}}(t)$): Outage is artificially declared, even though reliable transmission at rate R is possible.

Due to the definition of P^R, $\mathbb{P}(\mathcal{O}_{\text{ch}}(t))\leq\epsilon\ \forall t$, and the set $\{\mathcal{O}_{\text{ch}}(t)\}$ of events indexed by t are i.i.d. We choose $\{\mathcal{O}_{\text{a}}(t)\}$ such that $\mathcal{O}_{\text{x}}(t)=\mathcal{O}_{\text{ch}}(t)\cup\mathcal{O}_{\text{a}}(t)$ is i.i.d. as well, and

$$\mathbb{P}(\mathcal{O}_{\text{x}}(t))=\epsilon,\ \forall t$$

The dynamics of the *normalized* key queue[6] can therefore be modeled by

$$Q_M(t+1)=\min(M,Q_M(t)+R_s(t)-\mathbf{1}(\bar{\mathcal{O}}_{\text{enc}}(t))R) \quad (39)$$

Note that $Q_M(t)\geq 0\ \forall t$, due to the definition of $\mathcal{O}_{\text{key}}(t)$. Let $L^T(M)$ be the time average loss ratio over the first T blocks, for buffer size M, which is defined as the ratio of the amount of loss of key bits due to overflows, and the total amount of input key bits

$$L^T(M)=\frac{\sum_{t=1}^{T}\left(Q_M(t)+R_s(t)-\mathbf{1}(\bar{\mathcal{O}}_{\text{enc}}(t))R-M\right)^+}{\sum_{t=1}^{T}R_s(t)} \quad (40)$$

Then, we can see that $\forall T>0$,

$$(1-L^T(M))\sum_{t=1}^{T}R_s(t)=Q_M(T)+\sum_{t=1}^{T}R\mathbf{1}(\bar{\mathcal{O}}_{\text{enc}}(t)) \quad (41)$$

follows from (39), (40), and the fact that $Q_M(1)=0$.

Lemma 6: $Q_M(t)$ converges in distribution to an almost surely finite random variable.

The proof is provided in Appendix F-A. This implies that $\lim_{t\to\infty}\mathbb{P}(\mathcal{O}_{\text{enc}}(t))$ exists. Now, we provide our asymptotic result for the key loss ratio. We define the drift and variance of this process as

$$\begin{aligned}\mu_R&=\mathbb{E}[R_s(t)-R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))]\\&=\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]-R(1-\epsilon)\end{aligned} \quad (42)$$

and

$$\sigma_R^2=\mathbb{V}ar[R_s(t)-R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))]$$

respectively, where (42) follows from the definition of $\mathcal{O}_{\text{x}}(t)$.

Lemma 7: For any $M>0$, the key loss ratio satisfies the following asymptotic relationship

$$\lim_{R\searrow C_F(\epsilon)}\lim_{T\to\infty}L^T\left(M\frac{\sigma_R^2}{|\mu_R|}\right)\times\frac{2|\mu_R|\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]e^{\frac{-2R|\mu_R|}{\sigma_R^2}}}{\sigma_R^2}\leq e^{-2M} \quad (43)$$

The proof is provided in Appendix F-B.

Lemma 8: If $\lim_{t\to\infty}\mathbb{P}(\mathcal{O}_{\text{enc}}(t))=\epsilon'$, then ϵ' secrecy outage probability (7) is satisfied.

Proof: Find B such that $\mathbb{P}(\mathcal{O}_{\text{enc}}(t))=\epsilon'+\delta$ for any $t>B$. In 2-index time notation (s,b) with $t=sB+b$, it corresponds to $\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R))=\epsilon'+\delta,\ \forall(s,b):s\neq 1$. Then.

$$\begin{aligned}\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta))&\leq\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R))\\&\quad+\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)) &(44)\\&\leq\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)) &(45)\\&\leq\epsilon'+\delta &(46)\end{aligned}$$

[5]Note that, actual key buffer would be of size $NM(\epsilon,R)$ bits.

[6]Note that, the actual key queue process scales with N, i.e., $NQ_M(t)$ bits are in the key queue at block t.

Here, (44) follows from the union bound, and second term follows from (63) and (75) in Appendix A, which shows that there exists some packet size N large enough such that $\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) = 0$. Equation (46) implies that ϵ' secrecy outage probability (7) is satisfied. ∎

Let $\lim_{t\to\infty}\mathbb{P}(\mathcal{O}_{\text{enc}}(t)) = \epsilon'$. Since $\mathbb{P}(\mathcal{O}_{\text{x}}(t)) = \epsilon$ and $\mathcal{O}_{\text{enc}}(t) = \mathcal{O}_{\text{x}}(t) \cup \mathcal{O}_{\text{key}}(t)$, we have $\lim_{t\to\infty}\mathbb{P}(\mathcal{O}_{\text{key}}(t)) > 0$. This implies that $\lim_{T\to\infty}\frac{1}{T}Q_M(T) = 0$ (since otherwise, key outage probability would be zero), which, due to (41) implies

$$\begin{aligned}(1-\lim_{T\to\infty}L^T(M))\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))] &\\ = (1-\lim_{t\to\infty}\mathbb{P}(\mathcal{O}_{\text{enc}}(t)))R &\\ = (1-\epsilon')R &\end{aligned} \tag{47}$$

Here, due to the choice of power allocation policy P^R, we have $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))] = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}R_s(t)$. Plugging the result of Lemma 7 into (47), we obtain the required key buffer size to achieve ϵ' probability of secrecy outage

$$\lim_{R\searrow C_F(\epsilon)}\frac{M(\epsilon',R)-R}{\frac{\sigma_R^2}{2|\mu_R|}\log\left(\frac{\sigma_R^2}{2|\mu_R|(\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]-(1-\epsilon')R)}\right)} \le 1 \tag{48}$$

We know from (10) that ϵ and ϵ'-achievable secrecy capacities satisfy the conditions $(1-\epsilon')C_F(\epsilon') = \mathbb{E}[R_s(\mathbf{H},P^{C_F(\epsilon')}(\mathbf{H}))]$ and $(1-\epsilon)C_F(\epsilon) = \mathbb{E}[R_s(\mathbf{H},P^{C_F(\epsilon)}(\mathbf{H}))] = \mathbb{E}[R_s(\mathbf{H},P^*(\mathbf{H}))]$, respectively. By Lemma 3, we know that $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ is a continuous function of R, hence for any given $\epsilon' > \epsilon$, there exists an R such that $C_F(\epsilon) < R < C_F(\epsilon')$, and $\mathbb{E}[R_s(\mathbf{h},P^R(\mathbf{H}))] = (1-\frac{\epsilon+\epsilon'}{2})R$. Furthermore, as $\epsilon'\to\epsilon$, $C_F(\epsilon')\to C_F(\epsilon)$. Let us define a monotonically decreasing sequence $(\epsilon_1',\epsilon_2',\cdots)$, such that $\lim_{i\to\infty}\epsilon_i' = \epsilon$. For any $i\in\mathbb{N}$, find R_i such that $C_F(\epsilon) < R_i < C_F(\epsilon_i')$, and $\mathbb{E}[R_s(\mathbf{H},P^{R_i}(\mathbf{H}))] = (1-\frac{\epsilon+\epsilon_i'}{2})R_i$, therefore $\mu_{R_i} = (\epsilon-\epsilon')/(2R_i)$. From (48), we get

$$\lim_{i\to\infty}\frac{M(\epsilon_i',R_i)-R_i}{\kappa\left(\frac{\sigma_{R_i}^2}{R_i(\epsilon_i'-\epsilon)}\right)} \le 1 \tag{49}$$

We can see that (38) follows from (49), since as $i\to\infty$,

- $R_i\to C_F(\epsilon)$, where $C_F(\epsilon)<\infty$ converges as shown in (17), hence we can safely drop that term from the numerator, since the denominator diverges.
- $\epsilon_i'\to\epsilon$.
- $\sigma_{R_i}^2\to\sigma_{C_F(\epsilon)}^2$, where

$$\begin{aligned}\sigma_{C_F(\epsilon)}^2 &= \mathbb{V}ar[R_s(t)-C_F(\epsilon)\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))]\\ &\le \mathbb{V}ar[R_s(\mathbf{H},P^{C_F(\epsilon)}(\mathbf{H}))]-\epsilon(1-\epsilon)C_F(\epsilon).\end{aligned}$$

∎

VI. Numerical Results

In this section, we conduct simulations to illustrate our main results with two examples. In the first example, we analyze the relationship between ϵ-achievable secrecy capacity and average power. We assume that both the main channel and eavesdropper channel are characterized by Rayleigh fading, where the main channel and eavesdropper channel power gains follow exponential distribution with means 2 and 1, respectively. Since Rayleigh channel is non-invertible, maintaining a non-zero secrecy rate with zero secrecy outage probability is impossible. In Figure 3, we plot the ϵ-achievable secrecy capacity as a function of the average power, for $\epsilon = 0.02$ outage probability, for both full CSI and main CSI cases. It can be clearly observed from the figure that the gap between capacities under full CSI and main CSI vanishes as average power increases, which support the result of Theorem 3.

In the second example, we study the relationship between the buffer size, key loss ratio and the outage probability. We assume that both the main and eavesdropper channel gains follow a chi-square distribution of degree 2, but with means 2 and 1, respectively. We focus on the full CSI case, and consider the scheme described in Section V. We consider transmission at secrecy rate of R with the use of the power allocation policy P^R that solves the problem (25)-(27). For $\epsilon = 0.02$, and the average power constraint $P_{\text{avg}} = 1$, we plot the key loss ratio (40), as a function of buffer size M in Figure 4, for $R = C_F(\epsilon)$, $R = 1.01C_F(\epsilon)$ and $R = 1.02C_F(\epsilon)$, where $C_F(\epsilon)$ is the ϵ-achievable secrecy capacity. It is shown in Lemma 7 of Section V that expect the key loss ratio $L^T(M)$ decreases as R increases, which is observed in Figure 4. Finally, we study the relationship between the secrecy outage probability and the buffer size for a given rate. In Figure 5, we plot the secrecy outage probabilities, denoted as ϵ', as a function of buffer size M for the same encoder parameters. On the same graph, we also plot our asymptotic result given in Theorem 6, which provides an upper bound on the required buffer size to achieve ϵ' outage probability for rate $C_F(\epsilon)$, with the assumption that (38) is an equality for any ϵ'. We can see that, this theoretical result serves as an upper bound on the required buffer size when $\epsilon'-\epsilon$, which is the additional secrecy outages due to key buffer overflows, is very small. Another important observation from Figures 4 and 5 is that, for a fixed buffer size, although the key loss ratio decreases as R increases, secrecy outage probability increases. This is due to the fact that key bits are pulled from the key queue at a faster rate, hence the decrease in the key loss ratio does not compensate for the increase of the rate that key bits are pulled from the key queue, therefore the required buffer size to achieve same ϵ' is higher for larger values of R.

VII. Conclusions

This paper obtained sharp characterizations of the secrecy outage capacity of block flat fading channels under the assumption full and main CSI at the transmitter. In the two cases, our achievability scheme relies on opportunistically exchanging private keys between the legitimate nodes and using them later to secure the delay sensitive information. We further derive the optimal power control policy in each scenario revealing an interesting structure based by judicious time sharing between time sharing and the optimal strategy for the ergodic. Finally, we investigate the effect of key buffer overflow on the secrecy outage probability when the key buffer size is finite.

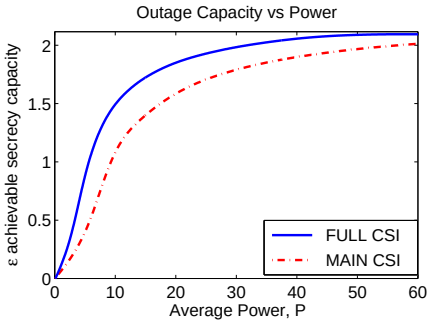

Fig. 3. The ϵ-achievable secrecy capacities as a function of average power, P_{avg}

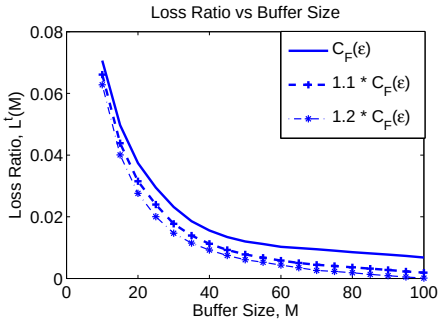

Fig. 4. Relationship between buffer size M, and key loss ratio $L^t(M)$

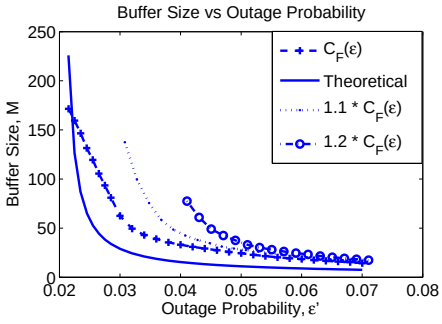

Fig. 5. Relationship between buffer size M, and outage probability ϵ'

APPENDIX A

A. Proof of Theorem 1

First, we prove the achievability. Let us fix $R < \mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]/(1-\epsilon)$, and consider a power allocation policy $P \in \mathcal{P}_F$, that satisfies the constraints (14),(15). We show that for any $\delta > 0$, there exist some B and N large enough such that the constraints in Definition 1 are satisfied, which implies that any $R < \mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]/(1-\epsilon)$ is an ϵ-achievable secrecy rate. The outage capacity is then found by maximizing $\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]/(1-\epsilon)$ based on constraints (10)-(14). For notational simplicity, we will use $R_s(s,b) \equiv R_s(\mathbf{H}(s,b), P(\mathbf{H}(s,b)))$ to denote the value of $R_s(\cdot)$ at block (s,b), and similarly define $R_m(s,b)$.

Our scheme, shown in Figure 6, utilizes secret key buffers at both the transmitter and legitimate receiver, where
i) At the end of every block (s,b), using privacy amplification, legitimate nodes (transmitter and receiver) generate $N(R_s(s,b)-\delta)$ bits of secret key from the transmitted signal in that particular block, and store it in their secret key buffers. We denote the generated secret key at the transmitter as $V(s,b)$, and at the receiver as $\hat{V}(s,b)$.
ii) At every block (s,b), $s \neq 1$, the transmitter pulls NR bits from its secret key buffer to secure the outage constrained message of size $H(W(s,b)) = NR$, using Vernam's one time pad. The receiver uses the same key to correctly decode the message. We denote the pulled key at the transmitter as $K(s,b)$, and at the receiver as $\hat{K}(s,b)$. Keys generated at $(s-1)$-st superblock are used only in the s-th superblock, and every generated key is only used *once*. When certain conditions are not met, this stage is skipped; the message $W(s,b)$ is not transmitted, and the keys are not pulled from the key queue. We call this particular event "encoder outage", and denote it as $\mathcal{O}_{\text{enc}}(s,b,R) \triangleq \mathcal{O}_{\text{ch}}(s,b,R) \cup \mathcal{O}_{\text{key}}(s,b,R) \cup \mathcal{O}_{\text{a}}(s,b,R)$, where

- Channel outage ($\mathcal{O}_{\text{ch}}(s,b,R)$): Channel is not suitable for reliable transmission at rate R, i.e., $R_m(s,b) < R$. Since P satisfies (14), due to the definition in (11), (12), for any (s,b) [7]
$$\mathbb{P}(\mathcal{O}_{\text{ch}}(s,b,R)) = \mathbb{P}\left(R_m(\mathbf{H}, P(\mathbf{H})) < \frac{\mathbb{E}[R_s(\mathbf{H}, P(\mathbf{H}))]}{1-\epsilon}\right) \leq \epsilon. \quad (50)$$
- Key outage ($\mathcal{O}_{\text{key}}(s,b,R)$): There are not enough key bits in the key queue to secure $W(s,b)$, i.e.,
$$\left(\sum_{b'=1}^{B} H(V(s-1,b')) - \sum_{b'=1}^{b} H(K(s,b'))\right) < 0$$
- Artificial outage ($\mathcal{O}_{\text{a}}(s,b,R)$): The transmitter declares 'outage', even though reliable secure transmission of $W(s,b)$ is possible. This is introduced to control the key queue dynamics and bound the probability of key outages, which is covered in the secrecy outage analysis. By definition, the events $\{\mathcal{O}_{\text{a}}(s,b,R)\}_{s=1,b=1}^{S,B}$ are mutually independent, they are also independent of other random variables, and satisfy the equality
$$\mathbb{P}(\mathcal{O}_{\text{ch}}(s,b,R) \cup \mathcal{O}_{\text{a}}(s,b,R)) = \epsilon. \quad (51)$$
for any (s,b).

Note that, due to our assumption that keys generated in $(s-1)$-st superblock are used in s-th superblock, all the blocks the first superblock $(s=1)$ observe key outages, therefore secrecy outages, yet it does not violate the constraints in Definition 1. Also note that, we will show that for any $\delta > 0$, there exist S, B and N are large enough such that the events $\mathcal{O}_{\text{enc}}(s,b,R)$ and $\mathcal{O}_{\text{inf}}(s,b,R,\delta)$ are equivalent.

Encoding:
Our random coding arguments rely on an ensemble of codebooks generated according to a zero mean Gaussian distribution with variance $P(s,b)$ [8].
1) When $\mathcal{O}_{\text{enc}}(s,b,R)$ does not occur, the message is secured with the secret key bits pulled from the key queue, using one time pad[9]
$$W_{sec}(s,b) = W(s,b) \oplus K(s,b) \quad (52)$$

[7]Here, we interchangeably use $R_m(s,b) \equiv R_m(\mathbf{H}(s,b), P(\mathbf{H}(s,b)))$, and due to stationarity of P, drop index (s,b).

[8]Note that, it is also possible to use a finite number of codebooks by partitioning the set $\{\mathbf{h}\}$ of channel gains, and using a different Gaussian codebook for every partition [3].

[9]We assume that both the message and the key are converted to binary form in this process.

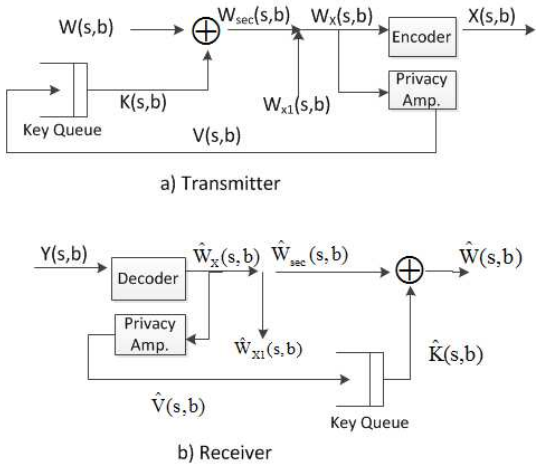

Fig. 6. The capacity achieving scheme, transmitter and receiver operation when $\mathcal{O}_{\text{enc}}(s,b,R)$ does not occur.

Clearly, $W_{sec}(s,b) \in \mathcal{W}_{sec} = \{1,\cdots,2^{NR}\}$. Furthermore, let $\{W_{x1}(s,b)\}_{s=1,b=1}^{S,B}$ denote an i.i.d. sequence where $\{W_{x1}(s,b)\} \in \{1,\cdots,2^{N(R_m(s,b)-R-\delta)}\}$ is uniformly distributed. The encoder forms $W_X(s,b) = [W_{sec}(s,b)\ W_{x1}(s,b)]$ by concatenation, and transmits the codeword $\mathbf{X}(s,b)$ indexed by $W_X(s,b)$ over the channel.

2) When $\mathcal{O}_{\text{enc}}(s,b,R)$ occurs, $W(s,b)$ is not transmitted. Let $\{W_{x2}(s,b)\}_{s=1,b=1}^{S,B}$ denote an i.i.d. sequence where $\{W_{x2}(s,b)\} \in \{1,\cdots,2^{N(R_m(s,b)-\delta)}\}$ is uniformly distributed. The encoder forms $W_X(s,b) = [W_{x2}(s,b)]$, and transmits the codeword $\mathbf{X}(s,b)$ indexed by $W_X(s,b)$ over the channel.

The reason for transmitting $W_{x1}(s,b)$ and $W_{x2}(s,b)$ is to confuse the eavesdropper to the fullest extent in the privacy amplification process.

Decoding:

The receiver finds the jointly typical $(\hat{W}_X(s,b), \mathbf{Y}(s,b))$ pair, where

1) $\hat{W}_X(s,b) = [\hat{W}_{sec}(s,b)\ \hat{W}_{x1}(s,b)]$ when $\mathcal{O}_{\text{enc}}(s,b,R)$ does not occur.
2) $\hat{W}_X(s,b) = [\hat{W}_{x2}(s,b)]$ when $\mathcal{O}_{\text{enc}}(s,b,R)$ occurs.

Define the error events

$$E_1(s,b) = \left\{\hat{W}_X(s,b) \neq W_X(s,b))\right\}$$
$$E_2(s,b,\delta) = \left\{\frac{1}{N}\|X(s,b)\|^2 > P(s,b)+\delta\right\}$$

Note that, the main channel at slot (s,b) can be viewed as an Additive White Gaussian Noise (AWGN) channel with channel gain $\mathbf{H}(s,b)$, which has instantaneous capacity of $R_m(s,b) \equiv R_m(\mathbf{H}, P(\mathbf{H}))$ [9]. The encoding rate (rate of $W_X(s,b)$) is equal to $R_m(s,b)-\delta$, which is below the instantaneous main channel capacity. Therefore, random coding arguments guarantee us that $\forall B>0$, $\exists N_1(B)>0$ such that $\forall N \geq N_1(B)$, $\mathbb{P}(E_1(s,b)) \leq \frac{\delta}{3B}$ and $\mathbb{P}(E_2(s,b,\delta)) \leq \frac{\delta}{3}$.

Privacy Amplification: At the end of every block (s,b), the transmitter and receiver generate secret key bits, by applying a universal hash[10] function on the exchanged signals in that particular block. First, we provide the definition of a universal hash function.

Definition 3: ([11]) A class G of functions $\mathcal{A} \to \mathcal{B}$ is universal, if for any $x_1 \neq x_2$ in $\mathcal{A}$, the probability that $g(x_1) = g(x_2)$ is at most $\frac{1}{\mathcal{B}}$ when g is chosen at random from G according to a uniform distribution.

Lemma 9: For any $S>0$, $B>0$, there exists $N_2(S,B) > 0$ such that, $\forall N \geq N_2(S,B)$, and for any block (s,b), the transmitter and receiver generate secret key bits $V(s,b) = G(W_X(s,b))$ and $\hat{V}(s,b) = G(\hat{W}_X(s,b))$ respectively, such that $V(s,b) = \hat{V}(s,b)$ if the error event $E_1(s,b)$ does not occur, and

$$H(V(s,b)) = N(R_s(s,b)-\delta) \quad (53)$$
$$\frac{1}{N} I(V(s,b); \mathbf{Z}(s,b), \mathbf{h}(s,b), G) \leq \frac{\delta}{SB} \quad (54)$$

The proof follows the approach of [12], which applies privacy amplification to Gaussian channels. We provide it in Appendix A-B.

Now, we will show that for this scheme, the error constraint in (6), and the secrecy outage constraint in (7) is satisfied. To simplify notation, we will frequently use $W(s,:) = \{W(s,b)\}_{b=1}^{B}$. Note that, for any s, the markov chain in Figure 7 is satisfied, which could be observed from Figure 6. These markov relations will be repeatedly used in secrecy outage analysis.

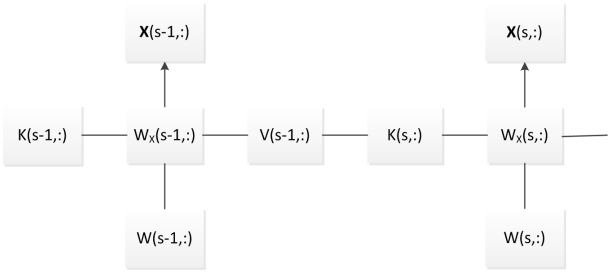

Fig. 7. Markov chains

Error Analysis: The probability of error event in (6) can be bounded as

$$\begin{aligned}\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)) &\\ \leq \mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) &\\ +\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)|\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)) & \quad (55)\\ = \mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) & \quad (56)\end{aligned}$$

where (56) follows since

$$\begin{aligned}\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)|\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)) &\\ \leq \mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)|\bar{\mathcal{O}}_{\text{inf}}(s,b,R,\delta)) = 0 & \quad (57)\end{aligned}$$

due to the fact that information outage ($\mathcal{O}_{\text{inf}}(s,b,R,\delta)$) does not occur, then $\frac{1}{N}H(W(s,b)|\mathbf{Y}^{s,b}, \mathbf{h}^{s,b}) > R-\delta$ which eliminates the possibility of an encoder outage ($\mathcal{O}_{\text{enc}}(s,b,R)$).

[10]Privacy amplification can also be performed using extractor functions. In [10], it is shown that in fading Gaussian channels, same key rates can be achieved by using extractor functions, as compared to universal hash functions.

For any $N \geq \max(N_1(B), N_2(S,B))$, and (s,b) such that $s>1$, (56) can be bounded as

$$\begin{aligned}
&\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) \\
&\leq \mathbb{P}(W(s,b)\neq \hat{W}(s,b)) + \mathbb{P}\left(\frac{\|\mathbf{X}(s,b)\|^2}{N} > P(s,b)+\delta\right) && (58)\\
&= \mathbb{P}(W_{sec}(s,b)\oplus K(s,b) \neq \hat{W}_{sec}(s,b)\oplus \hat{K}(s,b)) \\
&\quad + \mathbb{P}\left(\frac{\mathbf{X}(s,b)\|^2}{N}\| > P(s,b)+\delta\right) && (59)\\
&\leq \mathbb{P}(W_{sec}(s,b)\neq \hat{W}_{sec}(s,b)) + \mathbb{P}(K(s,b)\neq \hat{K}(s,b)) \\
&\quad + \mathbb{P}\left(\frac{\|\mathbf{X}(s,b)\|^2}{N} > P(s,b)+\delta\right) && (60)
\end{aligned}$$

where (58) follows from (2), and the union bound, (59) follows from the fact that when $\mathcal{O}_{\text{enc}}(s,b,R)$ does not occur, $W_{sec}(s,b) = W(s,b)\oplus K(s,b)$, and (60) follows from the union bound. The first term of (60) can be bounded as $\mathbb{P}(W_{sec}(s,b)\neq \hat{W}_{sec}(s,b)) \leq \frac{\delta}{3B}$ due to definition of $E_1(s,b)$, and the choice of N. Similarly, the third term can be bounded as $\mathbb{P}\left(\frac{1}{N}\|\mathbf{X}(s,b)\|^2 > P(s,b)+\delta\right) \leq \delta/3$ due to definition of $E_2(s,b,\delta)$, and the choice of N. The second term can be bounded as

$$\begin{aligned}
&\mathbb{P}\big(K(s,b)\neq \hat{K}(s,b)\big) \\
&\overset{(a)}{\leq} 1 - \prod_{i=1}^{B}\mathbb{P}(V(s-1,i)=\hat{V}(s-1,i)) \\
&\leq \sum_{i=1}^{B}\mathbb{P}(V(s-1,i)\neq \hat{V}(s-1,i)) \\
&\leq \sum_{i=1}^{B}\mathbb{P}(E_1(s-1,i)) \overset{(b)}{\leq} B\frac{\delta}{3B}
\end{aligned}$$

where (a) follows from the fact that keys used in s-th superblock are generated in $(s-1)$-st superblock, and (b) follows due to the definition of $E_1(s,b)$. Therefore, the error constraint in (6) is satisfied.

Secrecy Outage Analysis: The following lemmas will be useful in the secrecy outage analysis.

Lemma 10: For any B, there exists some N_3 such that for $N \geq N_3(B)$, the events $\mathcal{O}_{\text{enc}}(s,b,R)$ and $\mathcal{O}_{\text{inf}}(s,b,R,\delta)$ coincide with probability 1, i.e.,

$$\begin{aligned}
&\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)|\bar{\mathcal{O}}_{\text{inf}}(s,b,R,\delta)) \\
&\quad + \mathbb{P}(\mathcal{O}_{\text{inf}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) = 0. && (61)
\end{aligned}$$

Proof of Lemma 10 is provided in Appendix A-B. Secrecy outage probability can be bounded above as

$$\begin{aligned}
\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta)) &= \mathbb{P}(\mathcal{O}_{\text{eq}}(s,b,R,\delta)\cup\mathcal{O}_{\text{inf}}(s,b,R,\delta)) \\
&= \mathbb{P}(\mathcal{O}_{\text{eq}}(s,b,R,\delta)\cup\mathcal{O}_{\text{enc}}(s,b,R)) && (62)\\
&\leq \mathbb{P}(\mathcal{O}_{\text{eq}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) \\
&\quad + \mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R)) && (63)
\end{aligned}$$

where the first equality follows from the definition of secrecy outage in (3), (62) follows from Lemma 10, and (63) follows from the union bound. Now, we upper bound the first term. Note that

$$\begin{aligned}
&\mathbb{P}(\mathcal{O}_{\text{eq}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) = \mathbb{P}\bigg(\frac{1}{N}H(W(s,b)|\mathbf{Z}^{S,B}, \\
&\quad \mathbf{h}^{S,B}, W^{S,B}\backslash W(s,b), G, \bar{\mathcal{O}}_{\text{enc}}(s,b,R)) \geq R-\delta\bigg) && (64)
\end{aligned}$$

by definition in (4), and the fact that the universal hash function G used is revealed to the eavesdropper, hence the entropy of $W(s,b)$ is conditioned on G as well. For $s>1$, we bound the equivocation as follows[11]

$$\begin{aligned}
&H(W(s,b)|\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b),G,\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) \\
&\geq H\big(W(s,b)|\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b),G, \\
&\quad \bar{\mathcal{O}}_{\text{enc}}(s,b,R),W_{sec}(s,b)\big) && (65)\\
&= H\big(K(s,b)|\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b),G, \\
&\quad \bar{\mathcal{O}}_{\text{enc}}(s,b,R),W_{sec}(s,b)\big) && (66)\\
&= H(K(s,b)) - I\big(K(s,b);\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b), \\
&\quad G,\bar{\mathcal{O}}_{\text{enc}}(s,b,R),W_{sec}(s,b)\big) \\
&\geq H(K(s,b)) - \\
&\quad I\big(V(s-1,:);\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b),G\big) && (67)\\
&= H(K(s,b)) - \\
&\quad I\big(V(s-1,:);\mathbf{Z}^{s-1,B},\mathbf{h}^{s-1,B},W^{s-1,B},G\big) && (68)\\
&= H(K(s,b)) - \sum_{i=0}^{s-2} I\bigg(V(s-1,:);\mathbf{Z}(s-1-i), \\
&\quad \mathbf{h}(s-1-i,:),W(s-1-i,:),G\big| \\
&\quad \big\{\mathbf{Z}(s-1-j),\mathbf{h}(s-1-j,:),W(s-1-j,:)\big\}_{j=0}^{i-1}\bigg) && (69)\\
&\geq H(K(s,b)) - \sum_{i=0}^{s-2} I\bigg(V(s-1,:), \\
&\quad \big\{\mathbf{Z}(s-1-j),\mathbf{h}(s-1-j,:),W(s-1-j,:)\big\}_{j=0}^{i-1}; \\
&\quad \mathbf{Z}(s-1-i),\mathbf{h}(s-1-i,:),W(s-1-i,:),G\bigg) && (70)\\
&\geq H(K(s,b)) - \sum_{i=0}^{s-2} I\bigg(V(s-1-i,:);\mathbf{Z}(s-1-i), \\
&\quad \mathbf{h}(s-1-i,:),W(s-1-i,:),G\bigg) && (71)\\
&\geq H(K(s,b)) - \sum_{i=1}^{S} I\big(V(i,:);\mathbf{Z}(i,:),\mathbf{h}(i,:),G\big) && (72)\\
&\geq H(K(s,b)) - \sum_{i=1}^{S}\sum_{j=1}^{B} I\big(V(i,j);\mathbf{Z}(i,j),\mathbf{h}(i,j),G\big) && (73)\\
&\geq N(R-\delta) && (74)
\end{aligned}$$

where (65) follows from the fact that conditioning reduces entropy, (66) follows due to $K(s,b) = W_{sec}(s,b)\oplus W(s,b)$, (67) since $K(s,b)$ is pulled from the key buffer, which contains the key bits $V(s-1,:)$ generated during superblock $s-1$,

[11] In (69) and (70), $\{\cdot\}_{j=0}^{i-1} = \emptyset$ for $i \leq 1$.

hence $H(K(s,b)|V(s-1,:))=0$. (68) follows due to the Markov relation in Figure 7, along with the Markov chain $W_X(s,b)\to\mathbf{X}(s,b)\to\mathbf{Z}(s,b)$, and the fact that $K(s,:)$ is independent of $W_{sec}(s,:)$. The independence of $K(s,:)$ from $W_{sec}(s,:)$ follows since $W(s,b)$ is perfectly compressed, i.e., is of size NR bits with entropy NR, the one time pad performs as a Vernam cipher. (69) follows due to the chain rule, (70) follows since for any random variables A,B,C, $I(A;B|C)\leq I(A,B;C)$, and (71) follows since

$$V(s,:),\{\mathbf{Z}(s-j),\mathbf{h}(s-j,:),W(s-j,:)\}_{j=0}^{i-1}\to V(s-i,:)\to\mathbf{Z}(s-i),\mathbf{h}(s-i,:),W(s-i,:)$$

forms a Markov chain for any s and i, which can be observed from Figure 7. (72) follows since for any i, $V(i,:)$ is independent of $W(i,:)$ due to the one time pad. Similarly, (73) follows since for any i,j,j' such that $j\neq j'$,

$$(\mathbf{Z}(i,j),\mathbf{h}(i,j))\to V(i,j)\to W_X(i,j)\to K(i,:)\to W_X(i,j')\to V(i,j')\to(\mathbf{Z}(i,j'),\mathbf{h}(i,j'))$$

and the fact that $W_X(i,j)$ and $W_X(i,j')$ is independent due to one time pad. Finally, (74) follows due to the privacy amplification result in Lemma 9. Then,

$$\mathbb{P}(\mathcal{O}_{\text{eq}}(s,b,R,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R))=0 \tag{75}$$

due to (64) and (74). Now, we bound the second term. By the union bound,

$$\begin{aligned}&\mathbb{P}(\mathcal{O}_{\text{enc}}(s,b,R))\\&\quad\leq\mathbb{P}(\mathcal{O}_{\text{key}}(s,b,R))+\mathbb{P}(\mathcal{O}_{\text{ch}}(s,b,R)\cup\mathcal{O}_{\text{a}}(s,b,R))\\&\quad=\mathbb{P}(\mathcal{O}_{\text{key}}(s,b,R))+\epsilon\end{aligned}\tag{76}$$

where (76) follows due to (51). For (s,b), $s\neq 1$

$$\begin{aligned}&\mathbb{P}(\mathcal{O}_{\text{key}}(s,b,R))\\&\quad=\mathbb{P}\left(\sum_{i=1}^{B}H(V(s-1,i))-\sum_{i=1}^{b}H(K(s,i))<0\right)\\&\quad=\mathbb{P}\Bigg(\sum_{i=1}^{B}N(R_s(s-1,i)-\delta)-\\&\qquad\sum_{i=1}^{b}NR\mathbf{1}(\bar{\mathcal{O}}_{\text{ch}}(s,i,R)\cap\bar{\mathcal{O}}_{\text{a}}(s,i,R))<0\Bigg)\\&\quad\leq\mathbb{P}\Bigg(\sum_{i=1}^{B}\big[R_s(s-1,i)-\delta-\\&\qquad R\mathbf{1}(\bar{\mathcal{O}}_{\text{ch}}(s,i,R)\cap\bar{\mathcal{O}}_{\text{a}}(s,i,R))\big]<0\Bigg)\end{aligned}\tag{77}$$

Note that, the terms $\{R_s(s-1,i)\}$ and $\{\mathbf{1}(\bar{\mathcal{O}}_{\text{ch}}(s,i,R)\cap\bar{\mathcal{O}}_{\text{a}}(s,i,R))\}$ in (77) are i.i.d. with respect to s and i, and are independent of each other. Therefore, the expression in (77) represents a random walk with expected drift $\mu=\mathbb{E}[R_s(\mathbf{H},P(\mathbf{H}))]-\delta-R(1-\epsilon)$ due to the definition of artificial outage $\mathcal{O}_{\text{a}}(s,b,R))$ in (51). For[12] $R\leq\frac{\mathbb{E}[R_s(\mathbf{H},P(\mathbf{H}))]-\delta}{1-\epsilon}$, $\mu>0$, hence by the law of large numbers, $\exists B_1>0$ such that $\forall B>B_1$, $\mathbb{P}(\mathcal{O}_{\text{key}}(s,b,R))<\delta$, $s\neq 1$. Therefore, for the choice $B=B_1$, $N=\max(N_1(B_1),N_2(S,B_1),N_3(B_1))$, $\mathbb{P}(\mathcal{O}_{\text{sec}}(s,b,R,\delta))\leq\epsilon+\delta$ due to (76), (63) and (75). Hence the secrecy outage constraint in (7) is satisfied. This concludes the achievability.

Now, we prove the converse. Consider a power allocation policy P, which satisfies the average power constraint in (1). Let R be an ϵ-achievable secrecy rate. We will show that $R<C_F(\epsilon)$. Let $\delta>0$. Then $\exists B_1,N_1$ such that $\forall B>B_1,N>N_1$

$$\begin{aligned}&\frac{1}{SBN}\sum_{s=1}^{S}\sum_{b=1}^{B}H(W(s,b)|\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b))\\&\qquad\geq\sum_{s=1}^{S}\sum_{b=1}^{B}\frac{1}{SB}(R-\delta)\mathbf{1}(\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta))&(78)\\&\qquad\geq(R-\delta)(1-\epsilon-\delta)&(79)\end{aligned}$$

where (78) follows directly from the definition of the event $\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)$, and (79) follows from applying the secrecy outage constraint (7), and the law of large numbers.

It follows from the converse proof of ergodic secrecy capacity [3], and law of large numbers that $\exists B_2,N_2$ such that for every S, $B>B_2$, and $N>N_2$, the time-average equivocation rate[13] is bounded as

$$\begin{aligned}&\frac{1}{SBN}\sum_{s=1}^{S}\sum_{b=1}^{B}H(W(s,b)|\mathbf{Z}^{S,B},\mathbf{h}^{S,B},W^{S,B}\backslash W(s,b))\\&\qquad\leq\frac{1}{SBN}H(W^{S,B}|\mathbf{Z}^{S,B},\mathbf{h}^{S,B})&(80)\\&\qquad\leq\limsup_{S,B\to\infty}\sum_{s=1}^{S}\sum_{b=1}^{B}\frac{1}{SB}R_s(s,b)+\delta.&(81)\end{aligned}$$

Also note that,

$$\begin{aligned}\bar{\mathcal{O}}_{\text{sec}}(s,b,R,\delta)&\supseteq\bar{\mathcal{O}}_{\text{inf}}(s,b,R,\delta)\\&\supseteq\left\{\frac{1}{N}I(W(s,b);\mathbf{Y}^{s,b},\mathbf{h}^{s,b})\geq R-\delta\right\}\\&\supseteq\left\{\frac{1}{N}I(\mathbf{X}(s,b);\mathbf{Y}(s,b))\geq R-\delta\right\}&(82)\\&\supseteq\{R_m(s,b)\geq R-\delta\}&(83)\end{aligned}$$

where (82) follows from the fact that $W(s,b)\to(\mathbf{X}(s,b),\mathbf{X}^{s,b-1})\to(\mathbf{Y}(s,b),\mathbf{X}^{s,b-1})\to\mathbf{Y}^{s,b}$ forms a Markov chain. From the converse of the coding theorem [13]; the mutual information expression in (82) is maximized when $\mathbf{X}(s,b)$ becomes a Gaussian random vector, and the supremum is the expression in (83).

From (78), (81) and (83), it follows that any ϵ-achievable rate R is bounded above as

$$\begin{aligned}&R\leq\limsup_{S,B\to\infty}\sum_{s=1}^{S}\sum_{b=1}^{B}\frac{1}{SB}R_s(s,b)/(1-\epsilon)&(84)\\&\text{subject to: }\mathbb{P}(R_m(s,b)\geq R)\leq\epsilon&(85)\\&\qquad\limsup_{S,B\to\infty}\frac{1}{SB}\sum_{s=1}^{S}\sum_{b=1}^{B}P(s,b)\leq P_{\text{avg}}&(86)\end{aligned}$$

[12]The reason for introducing artificial outages is to make sure that the expected drift is positive.

[13]For any reliable code that yields vanishing probability of error as $S,B,N\to\infty$.

Since $R_m(s,b)$ and $R_s(s,b)$ are both deterministic functions of the power $P(s,b)$ and instantaneous channel gains $\mathbf{h}(s,b)$, it follows that the power allocation function that maximizes the right hand side of (84)-(86) is a stationary function of instantaneous channel gains $\mathbf{h}(s,b)$. Interchanging the notations $P(s,b) \equiv P(\mathbf{h})$, $R_s(s,b) \equiv R_s(\mathbf{h}, P(\mathbf{h}))$ and $R_m(s,b) \equiv R_m(\mathbf{h}, P(\mathbf{h}))$, we can see that the right hand side of (84)-(86) becomes $C_F(\epsilon)$, which completes the proof.

B. Proofs of Lemmas used in Appendix A

Proof of Lemma 9: First, we introduce the information theoretic quantities required for the proof. For random variables A, B, define

- Renyi entropy of A as $\log \mathbb{E}[P_A(a)]$
- Min-entropy as A as $H_\infty(A) = \min_a \log\left(\frac{1}{\mathbb{P}_A(a)}\right)$.
- Conditional min-entropy of A given B as $H_\infty(A|B) = \inf_b H_\infty(A|B=b)$.
- δ-smooth min-entropy of A as $H_\infty^\delta(A) = \max_{A':\|\mathbb{P}_A - \mathbb{P}_{A'}\|<\delta} H_\infty(A')$.

Without loss of generality, we drop the block index (s,b) and R, and focus on the first block $(1,1)$, and assume the event $\mathcal{O}_{\text{enc}}$ does not occur. Let $W_X = [W_{sec}\ W_{x1}]$, with sample realization sequences denoted by w_x. Let $V = G(W_X)$, where G denotes a random universal hash function that maps W_X to to an r-bit binary message $V \in \{0,1\}^r$. Then, it is clear that if error event E_1 does not occur, $\hat{V} = V$ since $W_X = \hat{W}_X$, for any choice of G. To show that the security constraints (53)-(54) are satisfied, we cite the privacy amplification theorem, which is originally defined for discrete channels. For this purpose, we define a quantization function ϕ, with sensitivity parameter $\Delta = \sup_{\mathbf{z}} |\mathbf{z} - \phi(\mathbf{z})|$. Let $\mathbf{Z}^\Delta = \phi(\mathbf{Z})$ denote the quantized version of $\mathbf{Z}$. where $\mathbf{z}^\Delta$ denotes realization sequences. Then, by Theorem 3 of [11] there exists a universal function G such that [14]

$$H(G(W_X)|\mathbf{Z}^\Delta = \mathbf{z}^\Delta, G) \geq r - \frac{2^{r - R(W_X|\mathbf{Z}^\Delta = \mathbf{z}^\Delta)}}{\log 2}$$

Now, we relate this expression to the Shannon entropy of the message, conditioned on eavesdropper's actual received signal. Using the facts $H_\infty(W_X) \leq R(W_X)$ and $H_\infty(W_X|\mathbf{Z}^\Delta, G) \leq H_\infty(W_X|\mathbf{Z}^\Delta = \mathbf{z}^\Delta, G)$, it is easy to show that

$$H(G(W_X)|\mathbf{Z}^\Delta, G) \geq r - \frac{2^{r - H_\infty(W_X|\mathbf{Z}^\Delta)}}{\log 2}$$

Then, due to the asymptotic relationship between continuous random variables and their quantized versions [13], there exists a quantization function ϕ such that Δ is small enough, and

$$\begin{aligned} H(G(W_X)|G, \mathbf{Z}) &\geq H(G(W_X)|G, \mathbf{Z}^\Delta) - \frac{\delta}{2SB} \\ &\geq r - \frac{2^{r - H_\infty(W_X|\mathbf{Z}^\Delta)}}{\log 2} - \frac{\delta}{2SB} \end{aligned} \tag{87}$$

[14]We omit $\mathbf{h}^{S,B}$ in the following parts of the proof of Lemma 9 for notational simplicity.

are satisfied. To relate min-entropy to Shannon entropy, we use the result of Theorem 1 of [12]; $\forall \delta' > 0$, $\exists$ a block length N' such that $\forall N > N'$,

$$\frac{1}{N} H(\mathbf{X}^\Delta|\mathbf{Z}^\Delta) \leq \frac{1}{N} H_\infty^{\delta'}(\mathbf{X}^\Delta|\mathbf{Z}^\Delta) + \delta/(SB) \tag{88}$$

Now, we proceed as follows,

$$\begin{aligned} H_\infty(W_X|\mathbf{Z}^\Delta) &= \lim_{\delta' \to 0} H_\infty^{\delta'}(W|\mathbf{Z}^\Delta) \\ &\geq H(W_X) - I(W_X; \mathbf{Z}^\Delta) - N\delta/(SB) && (89) \\ &\geq H(W_X) - I(\mathbf{X}; \mathbf{Z}) - N\delta/(SB) && (90) \\ &= NR_s - N\delta/(SB) && (91) \end{aligned}$$

where (89) follows from (88), and the appropriate choice of N'. (90) follows from the fact that $W_X \to \mathbf{X} \to \mathbf{Z} \to \mathbf{Z}^\Delta$ forms a Markov chain. (91) follows from the fact that $H(W_X) = N(R_m - \delta)$, and similarly $I(\mathbf{X}; \mathbf{Z}) \leq N(R_m - R_s - \delta)$, which is the eavesdropper's maximum achievable rate. Let $N'' = \frac{SB}{SB-1} \log\left(\frac{\delta \log(2)}{2SB}\right)$. For the choice of $H(V) = r = N(R_s - \delta)$, and $N \geq \max(N', N'')$, we get

$$\begin{aligned} I(V; G, \mathbf{Z}) &= H(G(W_X)) - H(G(W_X)|\mathbf{Z}, G) \\ &\leq \frac{2^{-N(B-1)/B}}{\log 2} + \frac{\delta}{2SB} && (92) \\ &\leq \frac{\delta}{SB} && (93) \end{aligned}$$

where (92) follows from (87), (91), and the fact that $V = G(W_X)$. (93) follows due to the choice of N''. Hence, for $N \geq \max(N', N'')$, the constraints (53), (54) are satisfied.

Proof of Lemma 10: The probability of the first term is 0 due to (57). For the second term, note that

$$\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R)) < \delta.$$

From Fano's inequality, we get

$$\begin{aligned} \frac{1}{N} H(W(s,b)|\mathbf{Y}^{s,b}, \mathbf{h}^{s,b}, \bar{\mathcal{O}}_{\text{enc}}(s,b,R)) \\ \leq \frac{1}{N} H(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R))) + \\ NR\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R))) \\ \leq \delta \end{aligned} \tag{94}$$

In the error analysis, it is shown that $\mathbb{P}(E(s,b,\delta)|\bar{\mathcal{O}}_{\text{enc}}(s,b,R))$ can be made arbitrarily small with increasing block length N, which shows that there exists N_3 such that for $N = \max(N_1(B), N_2(S,B), N_3(B))$, (94) holds, which proves that the probability of the second term is also 0.

Appendix B
Proof of Theorem 2

The proof is very similar to the proof for full CSI, hence we only point out the differences. For full CSI, key generation occurs at the end of every *block*, using privacy amplification. Due to lack of eavesdropper channel state at the legitimate nodes, this is no longer possible. However, as shown in [3], it is still possible to generate secret key bits over a *superblock*. The following lemma replaces Lemma 9 in the full CSI case.

Note that, we will use the notation $W(s,:)=\{W(s,b)\}_{b=1}^{B}$ for simplicity.

Lemma 11: Let $W_X(s,b)$ be defined as in full CSI case, where

$$W_X(s,b)= \begin{cases} [W_{sec}(s,b)\ W_{x1}(s,b)], \text{ if } \mathcal{O}_{\text{enc}}(s,b,R) \text{ does not occur} \\ [W_{x2}(s,b)], \text{ if } \mathcal{O}_{\text{enc}}(s,b,R) \text{ occurs} \end{cases}$$

There exists $N_2>0, B_1>0$ such that, $\forall N>N_2, B>B_1$, and for any superblock s, the transmitter and the receiver generates secret key bits $V(s)=G(W_X(s,:))$ and $\hat{V}(s)=G([\hat{W}_X(s,:))$ respectively, such that $V(s)=\hat{V}(s)$ if none of the error events $E_1(s,:)$ occur in superblock s, and

$$H(V(s))=NB\big(\mathbb{E}[R_s(\mathbf{H},P(H_m))]-\delta\big) \tag{95}$$

$$\frac{1}{N}I(V(s);\mathbf{Z}(s,:),\mathbf{h}(s,:),G)\leq\frac{\delta}{S} \tag{96}$$

The proof is very similar to the proof of Lemma 9, and is omitted here. Following the same error and secrecy outage analysis in the full CSI case, we can see that any rate $R<C_M(\epsilon)$ is achievable. The converse proof is also the same as in full CSI case, and is omitted here.

Appendix C
Proofs of Results in Section IV-A

A. Proof of Lemma 1

The parameter $R_{\max}$ is the maximum value for which the problem (25)-(27) has a solution; hence the average power constraint (26) is active. Moreover, the outage constraint (27) is also active, and due to the fact that $R_m(\mathbf{h},P)$ is a concave increasing function of P, we have $\mathbb{P}(R_m(\mathbf{H},P^{R_{\max}}(\mathbf{H}))=R_{\max})=(1-\epsilon)$, since otherwise one can further increase $R_{\max}$ to find a power allocation function that satisfies the equality. Since for a given $\mathbf{h}=[h_m\ h_e]$, the power allocation function that yields $R_{\max}$ is $P_{\text{inv}}(h_m,R_{\max})$, we have

$$P_{\text{avg}}=\int_{\mathbf{h}\in\mathcal{K}}P_{\text{inv}}(h_m,R_{\max})f(\mathbf{h})d\mathbf{h}$$

where $\mathcal{K}$ the set of channel gains for which the system operates at rate $R_{\max}$, and $\mathbb{P}(\mathbf{H}\in\mathcal{K})=(1-\epsilon)$. The set $\mathcal{K}$ contains channel gains $\mathbf{h}$ for which $P_{\text{inv}}(h_m,R_{\max})$ takes minimum values, so that the average power constraint is satisfied for the maximum possible R. Since $P_{\text{inv}}(\mathbf{h},R)=\frac{2^R-1}{h_m}$ is a decreasing function of h_m, one can see that the choice of $\mathcal{K}$ that yields $R_{\max}$ is $\mathcal{K}=\{\mathbf{h}:h_m\geq c\}$. Since the probability density function of $\mathbf{H}$ is well defined, $\mathbb{P}(H_m=0)=0$, hence $c>0$, which, along with $P_{\text{avg}}>0$, implies that $R_{\max}>0$.

B. Proof of Lemma 3

Let $R_{\max}>R>R'>0$. Then, any policy P that satisfies $\mathbb{P}(R_m(\mathbf{H},P(\mathbf{H}))<R)\leq\epsilon$, would also satisfy $\mathbb{P}(R_m(\mathbf{H},P(\mathbf{H}))<R')\leq\epsilon$. So, the set of power allocation functions that satisfy (27) shrinks as R increases, hence $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ is a non-increasing function of R. Now, we prove that $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ is continuous. From Lemma 2, we know that

$$\begin{aligned}P^R(\mathbf{h})&=P_{\text{wf}}(\mathbf{h},\lambda_R)+\\&\quad\mathbf{1}(\mathbf{h}\in\mathcal{G}(\lambda_R,k_R))(P_{\text{inv}}(h_m,R)-P_{\text{wf}}(\mathbf{h},\lambda_R))^+\\P^{R'}(\mathbf{h})&=P_{\text{wf}}(\mathbf{h},\lambda_{R'})+\\&\quad\mathbf{1}(\mathbf{h}\in\mathcal{G}(\lambda_{R'},k_{R'}))(P_{\text{inv}}(h_m,R')-P_{\text{wf}}(\mathbf{h},\lambda_{R'}))^+\end{aligned}$$

where (λ_R,k_R) and $(\lambda_{R'},k_{R'})$ are constants that satisfy (23) and (24) with equality with respect to parameters R and R', respectively. Due to the fact that the functions $P_{\text{inv}}(h_m,R)$ is continuous and monotone increasing with respect to R, $P_{\text{wf}}(\mathbf{h},\lambda)$ is continuous and monotone increasing with respect to λ, and the fact that integration preserves continuity, for any $\delta>0$ such that $R>R'>R-\delta$, we can find $\gamma>0$ such that

1) $\gamma>\lambda_R-\lambda_{R'}>0$
2) $\gamma>P_{\text{wf}}(\mathbf{h},\lambda_{R'})-P_{\text{wf}}(\mathbf{h},\lambda_R)>0,\ \forall\mathbf{h}$
3) $\mathbb{P}(\mathbf{H}\in\mathcal{G}(\lambda_{R'},k_{R'})\backslash\mathcal{G}(\lambda_R,k_R))<\gamma$
4) $|k_{R'}-k_R|<\gamma$
5) $\gamma>P_{\text{inv}}(h_m,R)-P_{\text{inv}}(h_m,R')>0,\ \forall\mathbf{h}$

Finally, due to the fact that $R_s(\mathbf{h},P)$ is a continuous and monotone increasing function of power P, and items $1-2$, we conclude that $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ is continuous.

C. Proof of Lemma 4

If $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]|_{R=0}=0$, then the unique solution of $R=\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]/(1-\epsilon)$ is $R=0$. So, consider $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]|_{R=0}=0$. It is easy to see that, $\frac{\mathbb{E}[R_s(\mathbf{H},P^{R_{\max}}(\mathbf{H}))]}{R_{\max}}\leq(1-\epsilon)$, since

$$\begin{aligned}\mathbb{E}[R_s(\mathbf{H},P^{R_{\max}}(\mathbf{H}))]&=\int_{h_m\geq c}R_s(\mathbf{h},P(\mathbf{h}))f(\mathbf{h})d\mathbf{h}\\&\leq R_m(\mathbf{h},P(\mathbf{h}))(1-\epsilon)\\&=R_{\max}(1-\epsilon)\end{aligned}$$

follows from definition of parameter c, and the inequality $R_s(\mathbf{h},P(\mathbf{h}))\leq R_m(\mathbf{h},P(\mathbf{h}))$. Combining the facts that, the function $\frac{\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]}{R}$ is continuous and strictly decreasing on $(0,R_{\max}]$, $\lim_{R\to0^+}\frac{\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]}{R}=\infty$ and $\frac{\mathbb{E}[R_s(\mathbf{H},P^{R_{\max}}(\mathbf{H}))]}{R_{\max}}\leq(1-\epsilon)$, by the intermediate value theorem, there exists a unique $R>0$, which satisfies $R=\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]/(1-\epsilon)$.

Appendix D
Proof of Lemma 2

We use Lagrangian optimization approach to find P^R. We can express $\mathbb{E}[R_s(\mathbf{H},P^R(\mathbf{H}))]$ given in (25)-(27) as

$$\begin{aligned}&\max_{P,\mathcal{G}}J(P)\\\text{s.t }&R_m(\mathbf{h},P(\mathbf{h}))\geq R,\quad\forall\mathbf{h}\in\mathcal{G}\\&\mathbb{P}(\mathbf{H}\in\mathcal{G})=1-\epsilon\end{aligned}\tag{97}$$

where the Lagrangian $J(P)$ is given by the equation[15]

$$J(P) = \int R_s(\mathbf{h}, P(\mathbf{h}))f(\mathbf{h})d\mathbf{h} - \lambda\left[\int P(\mathbf{h})f(\mathbf{h})d\mathbf{h} - P_{\text{avg}}\right] \quad (98)$$

Here, $\mathcal{G}$ is a set which consists of $\mathbf{h}$ for which $R_m(\mathbf{h}, P(\mathbf{h})) \geq R$ must be satisfied. We will show in this proof that it is of the form (20). This problem is identical to (25), since their constraint sets are identical. Hence solution of this problem would also yield P^R. In the following two-step approach, we proceed to find P^R. Let us fix $\lambda > 0$.

1) For any $\mathcal{G} \subseteq [0,\infty)\times[0,\infty)$, we find $P_{\mathcal{G}}$, which is defined as

$$P_{\mathcal{G}} = \arg\max_{P\in\mathcal{P}_F} J(P)$$
$$\text{s.t} \quad R_m(\mathbf{h}, P(\mathbf{h})) \geq R, \forall \mathbf{h} \in \mathcal{G} \quad (99)$$

2) Using the result of part 1, we find P^R, by finding the set $\mathcal{G}$ that maximizes $J(P)$, subject to a constraint $\mathbb{P}(\mathbf{H} \in \mathcal{G}) = 1-\epsilon$.

We start with step 1. Since both λ and R are fixed, therefore we drop them from $P_{\text{inv}}(\cdot)$ and $P_{\text{wf}}(\cdot)$, in the following parts to simplify the notation.

Lemma 12: If the problem (99) has a feasible solution, then it could be expressed as

$$P_{\mathcal{G}}(\mathbf{h}) = P_{\text{wf}}(\mathbf{h}) + [P_{\text{inv}}(\mathbf{h}) - P_{\text{wf}}(\mathbf{h})]^+ \mathbf{1}(\mathbf{h} \in \mathcal{G}) \quad (100)$$

where $P_{\text{wf}}(\mathbf{h})$ and $P_{\text{inv}}(\mathbf{h})$ are given in (19) and (18), respectively.

Proof: We will interchangeably use $\mathbf{h} = [h_m \ h_e]$. Due to (99), $R_m(\mathbf{h}, P(\mathbf{h})) = \log(1+P(\mathbf{h})h_m) \geq R$, $\forall \mathbf{h} \in \mathcal{G}$. Hence, there is a minimum power constraint for set $\mathcal{G}$, as

$$P(\mathbf{h}) \geq P_{\text{inv}}(\mathbf{h}) = \frac{2^R - 1}{h_m}, \forall \mathbf{h} \in \mathcal{G} \quad (101)$$

Define $\mathcal{K}$ as the set in which the minimum power constraint (101) is not active, i.e.,

$$\mathcal{K} = \{\mathbf{h} \in \mathcal{G} : P(\mathbf{h}) > P_{\text{inv}}(\mathbf{h})\} \cup \bar{\mathcal{G}}$$

where $\bar{\mathcal{G}}$ is complement of $\mathcal{G}$. First, we focus on the solution in the non-boundary set. Since the optimal solution must satisfy the Euler-Lagrange equations,

$$\frac{dJ(P)}{dP(\mathbf{h})} = 0, \mathbf{h} \in \mathcal{K}$$

For $\mathbf{h} \in \mathcal{K}$, we get the following condition

$$\frac{h_m}{1+h_m P(\mathbf{h})} - \frac{h_e}{1+h_e P(\mathbf{h})} - \lambda = 0$$

whose solution yields

$$P(\mathbf{h}) = \frac{1}{2}\left[\sqrt{\left(\frac{1}{h_e} - \frac{1}{h_m}\right)^2 + \frac{4}{\lambda}\left(\frac{1}{h_e} - \frac{1}{h_m}\right)} - \left(\frac{1}{h_e} + \frac{1}{h_m}\right)\right]$$

[15]Note that we leave the constraint (27) as is, and not include it in $J(P)$.

If for some $\mathbf{h} \in \mathcal{K}$, the value $P(\mathbf{h})$ is negative, then due to the concavity of $J(P)$ with respect to $P(\mathbf{h})$, the optimal value of $P(\mathbf{h})$ is zero [3]. Therefore, the solution yields

$$P(\mathbf{h}) = P_{\text{wf}}(\mathbf{h}), \quad \forall \mathbf{h} \in \mathcal{K} \quad (102)$$

Combining the result with the minimum power constraint inside set $\mathcal{G}$, the solution of (99) yields (100), which concludes the proof. ∎

Now, we find P^R. We proceed by further simplifying the Lagrangian in (98), for the case where $P = P_{\mathcal{G}}$, for a given $\mathcal{G}$ as follows.

$$\begin{aligned} J(P_{\mathcal{G}}) = & \int_{\mathbf{h}\in\mathcal{G}} [R_s(\mathbf{h}, P(\mathbf{h})) - \lambda P(\mathbf{h})]\, f(\mathbf{h})d\mathbf{h} \\ & + \int_{\mathbf{h}\notin\mathcal{G}} [R_s(\mathbf{h}, P(\mathbf{h})) - \lambda P(\mathbf{h})]\, f(\mathbf{h})d\mathbf{h} \\ = & \int [R_s(\mathbf{h}, P_{\text{wf}}(\mathbf{h})) - \lambda P_{\text{wf}}(\mathbf{h})]\, f(\mathbf{h})d\mathbf{h} \\ & + \int_{\mathcal{G}} \Big\{[R_s(\mathbf{h}, P_{\text{inv}}(\mathbf{h})) - R_s(\mathbf{h}, P_{\text{wf}}(\mathbf{h}))]^+ \\ & - \lambda\,[P_{\text{inv}}(\mathbf{h}) - P_{\text{wf}}(\mathbf{h})]^+\Big\} f(\mathbf{h})d\mathbf{h} \end{aligned} \quad (103)$$

After this simplification, the first term in (103) does not depend on $\mathcal{G}$. We conclude the proof by showing that $P^R = P_{\mathcal{G}^*}$ where the set $\mathcal{G}^*$ is defined as follows,

$$\mathcal{G}^* = \Big\{\mathbf{h} : [R_s(\mathbf{h}, P_{\text{inv}}(\mathbf{h})) - R_s(\mathbf{h}, P_{\text{wf}}(\mathbf{h}))]^+ - \lambda\,[P_{\text{inv}}(\mathbf{h}) - P_{\text{wf}}(\mathbf{h})]^+ \geq k\Big\} \quad (104)$$

where the parameter k is a constant that satisfies $\mathbb{P}(\mathbf{H} \in \mathcal{G}^*) = (1-\epsilon)$. We prove this by contradiction. First define $\xi(\mathbf{h}) = [R_s(\mathbf{h}, P_{\text{inv}}(\mathbf{h})) - R_s(\mathbf{h}, P_{\text{wf}}(\mathbf{h}))]^+ - \lambda\,[P_{\text{inv}}(\mathbf{h}) - P_{\text{wf}}(\mathbf{h})]^+$. Then, it follows from (103) that $\mathcal{G}^*$ is the set that maximize (103), so

$$\mathcal{G}^* = \arg\max_{\mathcal{G}} \int_{\mathcal{G}} \xi(\mathbf{h}) f(\mathbf{h})d\mathbf{h}$$

Assume that some other $\mathcal{G}' \neq \mathcal{G}^*$ is optimal, where $\mathbb{P}(\mathbf{H} \in \mathcal{G}') = 1-\epsilon$. However, we have

$$\begin{aligned} & J(P_{\mathcal{G}^*}) - J(P_{\mathcal{G}'}) \\ & = \int_{\mathcal{G}^*} \xi(\mathbf{h}) f(\mathbf{h})d\mathbf{h} - \int_{\mathcal{G}'} \xi(\mathbf{h}) f(\mathbf{h})d\mathbf{h} \\ & = \int_{\mathcal{G}^*\setminus\mathcal{G}'} \xi(\mathbf{h}) f(\mathbf{h})d\mathbf{h} - \int_{\mathcal{G}'\setminus\mathcal{G}^*} \xi(\mathbf{h}) f(\mathbf{h})d\mathbf{h} \\ & \geq 0 \end{aligned} \quad (105)$$

since

$$\int_{\mathcal{G}^*\setminus\mathcal{G}'} f(\mathbf{h})d\mathbf{h} = \int_{\mathcal{G}'\setminus\mathcal{G}^*} f(\mathbf{h})d\mathbf{h}$$

and

$$\xi(\mathbf{h})|_{\mathbf{h}\in\mathcal{G}^*} \geq \xi(\mathbf{h})|_{\mathbf{h}\in\mathcal{G}'}, \ \ \forall \mathbf{h}$$

by definition. This contradicts our assumption that $\mathcal{G}'$ is optimal. Note that, $\mathcal{G}^*$ is identical to (21). This concludes the proof.

APPENDIX E
PROOF OF LEMMA 5

The proof goes along similar lines as in Appendix D, so we skip the details here. We solve the problem for a fixed $\lambda > 0$. First, for any given $\mathcal{G} \in [0, \infty)$, we define the following problem, the solution of which yields $P_{\mathcal{G}}$.

$$P_{\mathcal{G}} = \arg \max_{P \in \mathcal{P}_M} J(P) \tag{106}$$

$$\text{subject to: } R_m(\mathbf{h}, P(h_m)) \geq R, \forall h_m \in \mathcal{G} \tag{107}$$

Lemma 13: If the problem (106) has a feasible solution, then it can be expressed as

$$P_{\mathcal{G}}(h_m) = P_w(h_m, \lambda) + \mathbf{1}(h_m \in \mathcal{G}) \left(P_{\text{inv}}(h_m, R) - P_w(h_m, \lambda)\right)^+ \tag{108}$$

Proof: The proof uses the same approach as in proof of Lemma 12. We define the set $\mathcal{K}$ such that for any $h_m \in \mathcal{K}$, the minimum rate constraint in (107) is not active. Since the optimal solution must satisfy the Euler Lagrange equations, we have

$$\frac{dJ(P(h_m))}{dP(h_m)} = 0, \ h_m \in \mathcal{K}$$

If we solve the equation for any given h_m, we get

$$\frac{h_m \mathbb{P}(H_e \leq h_m)}{1 + h_m P(h_m)} - \int_0^{h_m} \left(\frac{h_e}{1 + h_e P(h_m)}\right) f(h_e) dh_e = \lambda$$

If the power allocation function that solves the equation is negative, then by the convexity of the objective function [3], the optimal value of $P(h_m)$ is 0. Hence, we get $P_w(\mathbf{h}, \lambda)$ as the resulting power allocation function. Whenever the minimum rate constraint (37) is active, we get the channel inversion power allocation function, $P_{\text{inv}}(\mathbf{h}, R)$. ∎

Now, using Lemma 13, we solve the following problem,

$$\max_{P, \mathcal{G}} J(P) \tag{109}$$
$$\text{s.t } R_m(\mathbf{h}, P(\mathbf{h})) \geq R, \quad \forall \mathbf{h} \in \mathcal{G}$$
$$\mathbb{P}(H_m \in \mathcal{G}) = 1 - \epsilon$$

the solution of which yields P^R. Lemma 13 proves that the solution is a time-sharing between policies P_w and P_{inv}. Now, we find the optimal $\mathcal{G}$.

Lemma 14: The solution of (109) is of the form (108), with the set $\mathcal{G}^* = [c, \infty)$, where c is a constant which solves $\mathbb{P}(H_m \geq c) = 1 - \epsilon$.

Proof: Let $P_{\mathcal{G}^*}$ and $P_{\mathcal{G}'}$ be the power allocation functions that are solutions of (108) given the sets $\mathcal{G}^*$ and $\mathcal{G}'$, respectively. We show that, any choice of $\mathcal{G}' \neq \mathcal{G}^*$, such that $\mathbb{P}(H_m \in \mathcal{G}') = 1 - \epsilon$ is suboptimal, i.e.,

$$J(P_{\mathcal{G}^*}) - J(P_{\mathcal{G}'}) \geq 0$$

We continue as follows. To simplify notation, for $\mathbf{h} = [h_m \ h_e]$, let us denote

$$\xi_R(\mathbf{h}) = \left[R_s(\mathbf{h}, P_{\text{inv}}(h_m, R)) - R_s(\mathbf{h}, P_w(h_m, \lambda))\right]^+$$
$$\xi_P(h_m) = \left[P_{\text{inv}}(h_m, R) - P_{\text{wf}}(h_m, \lambda)\right]^+$$

Then,

$$J(P_{\mathcal{G}^*}) - J(P_{\mathcal{G}'}) = \int_{h_e} \left\{ \int_{h_m \in \mathcal{G}^*} \left(\xi_R(\mathbf{h}) - \lambda \xi_P(h_m)\right) f(h_m) dh_m \right\} f(h_e) dh_e - \int_{h_e} \left\{ \int_{h_m \in \mathcal{G}'} \left(\xi_R(\mathbf{h}) - \lambda \xi_P(h_m)\right) f(h_m) dh_m \right\} f(h_e) dh_e$$

Note that, for any $h'_m \in \mathcal{G}^* \backslash \mathcal{G}'$ and $h''_m \in \mathcal{G}' \backslash \mathcal{G}^*$, we have $h'_m > h''_m$. Since $P_w(h'_m, \lambda) \geq P_w(h''_m, \lambda)$ and $P_{\text{inv}}(h'_m, \lambda) < P_{\text{inv}}(h''_m, \lambda)$, we have $\xi_P(h'_m) \leq \xi_P(h''_m)$. Since $R_s(\mathbf{h}, P)$ is a concave increasing function of P [3], and for $\frac{dP_w(\mathbf{h}, P)}{dP} = \lambda$ for any $\mathbf{h}$, we have

$$\xi_R([h'_m \ h_e]) - \lambda \xi_P(h'_m) \geq \xi_R([h''_m \ h_e]) + \lambda \xi_P(h''_m)$$

Combining this result with the packing arguments following (104) in Appendix D, we get

$$J(P_{\mathcal{G}^*}) - J(P_{\mathcal{G}'}) \geq 0$$

hence concluding the proof. Note that, this result can also be proved using the arguments of Section 4 in [14]. ∎

APPENDIX F
PROOFS IN SECTION V

A. Proof of Lemma 6

Due to Theorem 1.2 of Section VI in [15], it suffices to show that $Q_M(t)$ is a positive recurrent regenerative process. Note that $Q_M(t)$ is a Markov process with an uncountable state space $[0 \ M]$, since $Q_M(t)$ can be written as $Q_M(t+1) = \min(M, Q_M(t) + R_s(t) - \mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t) \cap \bar{\mathcal{O}}_{\text{key}}(t)))$ where $R_s(t)$ and $\bar{\mathcal{O}}_{\text{x}}(t)$ are i.i.d., and $\bar{\mathcal{O}}_{\text{key}}(t) = \left\{Q_M(t) + R_s(t) - R \geq 0\right\}$ depends only on $Q_M(t)$ and $R_s(t)$. Therefore, $Q_M(t+1)$ is independent of $\{Q_M(i)\}_{i=1}^{t-1}$ given $Q_M(t)$, hence Markovity follows. Now, we prove that $Q_M(t)$ is a recurrent regenerative process where regeneration occurs at times $t_1, t_2, \cdots$ such that $Q_M(t_i) = M$. A sufficient condition for this is to show that $Q_M(t)$ has an accessible atom [16].

Definition 4: An accessible atom M is a state that is hit with positive probability starting from any state, i.e., $\sum_{t=1}^{\infty} \mathbb{P}(Q_M(t) = M | Q_M(1) = i) > 0 \ \forall i$.

Lemma 15: $Q_M(t)$ has an accessible atom M.

Proof: Assume $Q_M(1) = i$, $i \in [0, M]$. Note that, $R_s(t)$ and $\mathcal{O}_{\text{x}}(t)$ are both i.i.d. Also note that, $\mathbb{P}\big(R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t)) > 0\big) > 0 \ \forall t$ [16]. Find $\gamma > 0$ such that $\mathbb{P}\big(R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t)) > \gamma\big) = \gamma \ \forall t$. Let $\eta_i = \lceil \frac{M-i}{\gamma} \rceil$. Then,

$$\mathbb{P}(Q_M(\eta_i + 1) = M | Q_M(1) = i) \geq \prod_{t=1}^{\eta_i} \mathbb{P}\bigg(R_s(t) + \mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t)) > \delta\bigg) \geq \gamma^{\eta_i} > 0$$

∎

Since $Q_M(t)$ is a regenerative process, we know that $t_2 - t_1, t_3 - t_2, \cdots$ are i.i.d. random variables. Define a random variable τ, with distribution identical to $t_{i+1} - t_i$. Now we

[16] Since considering otherwise would lead to the uninteresting scenario where there are no buffer overflows (since the key queue cannot grow), hence any buffer size $M > C_F(\epsilon')$ is sufficient to achieve ϵ' secrecy outage probability.

show that $Q_M(t)$ is positive recurrent, by showing $\mathbb{E}[\tau] < \infty$. Consider another recursion

$$Q'_M(t+1) = \min\left(M, Q'_M(t) + R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))\right)^+ \tag{110}$$

with $Q'_M(1) = Q_M(1)$. It is clear that $Q'_M(t)$ is also regenerative, where regeneration occurs at $\{t'_i\}$, where $Q_M(t'_i) = M$, and let τ' be equal in distribution to $t'_{i+1} - t'_i$.

Lemma 16:

$$\mathbb{E}[\tau] \leq \mathbb{E}[\tau']$$

Proof: It suffices to show that when $Q_M(t) \neq M$, $Q'_M(t) \leq Q_M(t)$. By induction, assuming $Q'_M(t) \leq Q_M(t)$, we need to verify that $Q'_M(t+1) \leq Q_M(t+1)$. Consider $Q_M(t+1) < M$. Then,

$$\begin{aligned} Q_M(t+1) &= \left(Q_M(t) + R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t) \cap \bar{\mathcal{O}}_{\text{key}}(t)\right)^+ \\ &\geq \left(Q_M(t) + R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))\right)^+ \\ &\geq \left(Q'_M(t) + R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))\right)^+ \\ &= Q'_M(t+1) \end{aligned}$$

■

Note that $Q'_M(t)$ is regenerative both at states 0 and M. Let $\mathbb{E}[\tau'_1]$ denote the expected time for the process $Q'_M(t)$ to hit 0 from M, and $\mathbb{E}[\tau'_2]$ denote the expected time to hit M from 0. Then,

$$\mathbb{E}[\tau'] \leq \mathbb{E}[\tau'_1] + \mathbb{E}[\tau'_2] \tag{111}$$

Since the key queue has a negative drift, i.e., $\mu_R = \mathbb{E}[R_s(\mathbf{H}, P^R(\mathbf{H})) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))] < 0$, it is clear that $\mathbb{E}[\tau'_1] < \infty$. Now, we show that $\mathbb{E}[\tau'_2] < \infty$. Following the approach of Lemma 15, find $\gamma > 0$ such that $\mathbb{P}\big(R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t)) > \gamma\big) = \gamma \ \forall t$. Let $\eta = \lceil M/\gamma \rceil$. Then, $\mathbb{P}(Q_M(\eta+1) = M | Q_M(1) = 0) \geq \gamma^\eta > 0$, and

$$\begin{aligned} \mathbb{E}[\tau'_2] &\leq \sum_{i=0}^{\infty} (\eta + i(\mathbb{E}[\tau'_1] + \eta))\gamma^\eta (1-\gamma^\eta)^i \\ &\leq \eta\gamma^\eta \sum_{i=0}^{\infty} (1-\gamma^\eta)^i + \sum_{i=0}^{\infty} (1-\gamma^\eta)^i i(\mathbb{E}[\tau'_1] + \gamma^\eta) \\ &< \infty \end{aligned}$$

The first inequality follows from the fact that with probability γ^η, $Q_M(t)$ hits M at η'th block and with probability $(1-\gamma^\eta)$, key queue goes back to state 0 at $(\mathbb{E}[\tau'_1] + \gamma^\eta)$'th block (on average). The last inequality follows from $0 < \gamma^\eta < 1$, and ratio test. This result, along with (111) and Lemma 16 concludes that $Q_M(t)$ is a positive recurrent regenerative process, which concludes the proof.

B. Proof of Lemma 7

We follow an indirect approach to prove the lemma. Let $\{Q(t)\}_{t=1}^{\infty}$ denote the key queue dynamics of the same system for the infinite buffer case ($M = \infty$). First, we use the heavy traffic results in [17] to calculate the overflow probability of the infinite buffer queue. Then, we relate the overflow probability of infinite buffer system to the loss ratio of the finite buffer queue. The dynamics of the infinite buffer queue is characterized by

$$Q(t+1) = Q(t) + R_s(t) - \mathbf{1}(\bar{\mathcal{O}}_{\text{enc}}(t))R \tag{112}$$

where $Q(1) = 0$. The heavy traffic results we will use are for queues that have a stationary distribution. Since it is not clear whether $Q(t)$ is stationary or not, we will upper bound $Q(t)$ by another stationary process $Q'(t)$, and the buffer overflow probability result we will get for $Q'(t)$ will serve as an upper bound for $Q(t)$.

Let $\{Q'(t)\}_{t\geq 1}$ be the process that satisfies the following recursion

$$Q'(t+1) = \left(Q'(t) + R_s(t) - R\mathbf{1}(\bar{\mathcal{O}}_{\text{x}}(t))\right)^+ \tag{113}$$

with $Q'(1) = 0$. First, we relate $Q'(t)$ to $Q(t)$.

Lemma 17:

$$Q(t) \leq Q'(t) + R, \ \forall t \tag{114}$$

Proof: Assuming $Q(t) \leq Q'(t) + R$, we need to show by induction that $Q(t+1) \leq Q'(t+1) + R$. There are two different scenarios.

1) If $Q'(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right) \geq 0$, then, using the facts $\bar{\mathcal{O}}_{\text{enc}}(t) = \bar{\mathcal{O}}_{\text{x}}(t) \cap \bar{\mathcal{O}}_{\text{key}}(t)$ and $Q'(t) \leq Q(t)$, we obtain

$$\begin{aligned} &Q(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{enc}}(t)\right) \\ &\quad \geq Q'(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right) \geq 0 \end{aligned}$$

which, using the described key queue recursions in (112), implies

$$Q(t+1) = Q(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right) \tag{115}$$

Observe that, by (113),

$$Q'(t+1) = Q'(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right)$$

which, in conjunction with (115) and $Q(t) \leq Q'(t) + R$, yields $Q(t+1) \leq Q'(t+1) + R$.

2) If $Q'(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right) < 0$, then $Q'(t+1) = 0$. We further consider two cases. First, if $Q(t) + R_s(t) - R \geq 0$, then,

$$\begin{aligned} Q(t+1) &= \left(Q(t) + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right)\right)^+ \\ &\leq \left(Q'(t) + R + R_s(t) - R\mathbf{1}\left(\bar{\mathcal{O}}_{\text{x}}(t)\right)\right)^+ \\ &\leq Q'(t+1) + R = R \end{aligned} \tag{116}$$

Next, if $Q(t) + R_s(t) - R < 0$, then

$$Q(t+1) = Q(t) + R_s(t) < R = Q'(t+1) + R$$

which, combined with (116), yields $Q(t+1) \leq Q'(t+1) + R$.

■

Now, we show that $Q'(t)$ converges in distribution to an almost surely finite random variable Q'. First, we need to show that the expected drift of $Q'(t)$ is negative. It is clear from (113) that the expected drift of the process $Q'(t)$ is equal to $\mu_R = \mathbb{E}[R_s(\mathbf{H}, P^R(\mathbf{H}))] - R(1-\epsilon)$.

Lemma 18: For $R > C_F(\epsilon)$, we have $\mu_R < 0$, and μ_R is a continuous decreasing function of R.

Proof: From Lemma 3 in Section IV-A, we know that $\mathbb{E}[R_s(\mathbf{H}, P^R(\mathbf{H}))]$ is a non-increasing continuous function of R. Therefore, μ_R it is a continuous function of R. Furthermore, by definition of $C_F(\epsilon)$ in (10), $\mu_{C_F(\epsilon)} = 0$. Combining these two facts, we conclude that $\mu_R < 0$, for $R > C_F(\epsilon)$. ■

Lemma 19: There exists an almost surely finite random variable Q' such that, for all x,

$$\limsup_{t\to\infty} \mathbb{P}(Q(t) > x) \leq \mathbb{P}(Q' + R > x) \tag{117}$$

Proof: Combining Lemma 18 with the classic results by Loynes [18], we can see that $Q'(t)$ converges in distribution to an almost surely finite random variable Q' such that

$$\lim_{t\to\infty} \mathbb{P}(Q'(t) > x) = \mathbb{P}(Q' > x)$$

Using (114), we finish the proof of the lemma. ■

Now, we characterize the tail distribution of the key queue.

Lemma 20: For any given $M \geq 0$,

$$\lim_{R\searrow C_F(\epsilon)} \limsup_{t\to\infty} \mathbb{P}\left(\frac{|\mu_R|(Q(t)-R)}{\sigma_R^2} > M\right) \leq e^{-2M} \tag{118}$$

Proof: First, we prove that

$$\lim_{R\searrow C_F(\epsilon)} \mathbb{P}\left(\frac{|\mu_R|Q'}{\sigma_R^2} > y\right) = e^{-2y}, \tag{119}$$

which is based on the heavy traffic limit for queues developed in [17], see also Theorem 7.1 in [15]. In order to prove (119), we only need to verify the following three conditions: i) $\lim_{R\searrow C_F(\epsilon)} \mu_R = 0$; ii) $\lim_{R\searrow C_F(\epsilon)} \sigma_R^2 > 0$; and iii) the set $\left\{\left(R_s(\mathbf{H}, P^R) - R\mathbf{1}(\bar{\mathcal{O}}_{\mathrm{x}}(t))\right)^2\right\}$ of random variables indexed by R is uniformly integrable.

i) From Lemma 18, we obtain $\lim_{R\searrow C_F(\epsilon)} \mu_R = 0$.

ii) Since $R_s(\mathbf{H}, P^*(\mathbf{H})) - C_F(\epsilon)\bar{\mathcal{O}}_{\mathrm{x}}(t)$ is not a constant random variable, almost surely

$$\lim_{R\searrow C_F(\epsilon)} \sigma_R^2 = \mathbb{V}ar[R_s(\mathbf{H}, P^*(\mathbf{H})) - C_F(\epsilon)(\bar{\mathcal{O}}_{\mathrm{x}}(t))] > 0$$

iii) Note that, R lies on the interval $[0\ R_{\max}]$, where $R_{\max}$, defined in Lemma 1 then we have

$$\begin{aligned}
&\left(R_s(\mathbf{H}, P^R(\mathbf{H})) - R\mathbf{1}(\bar{\mathcal{O}}_{\mathrm{x}}(t))\right)^2 \\
&\quad = R_s(\mathbf{H}, P^R(\mathbf{H}))^2 \\
&\qquad - 2R_s(\mathbf{H}, P^R(\mathbf{H}))R\mathbf{1}\left(\bar{\mathcal{O}}_{\mathrm{x}}(t)\right) + R^2\mathbf{1}\left(\bar{\mathcal{O}}_{\mathrm{x}}(t)\right) \\
&\quad \leq R_s(\mathbf{H}, P^R(\mathbf{H}))^2 + R_{\max}^2
\end{aligned}$$

Since $R_s(\mathbf{h}, P)$ is a continuous function of P, and for any R on the interval $[0\ R_{\max}]$, $\lim_{c\to\infty} \mathbb{P}(P^R(\mathbf{H}) > c) = 0$, hence we can see that $\lim_{c\to\infty} \mathbb{P}(R_s(\mathbf{H}, P^R(\mathbf{H})) > c) = 0$. Therefore, this class of random variables is uniformly integrable. This completes the proof of (119). This result, in conjunction with Lemma 19 completes the proof. ■

Using Lemma 1 in [19], we relate the loss ratio of our finite buffer queue $Q_M(t)$ to the overflow probability of the infinite buffer queue $Q(t)$ as follows

$$\begin{aligned}
&\mathbb{E}[R_s(\mathbf{H}, P^R(\mathbf{H}))] \limsup_{T\to\infty} L^T(M) \\
&\qquad \leq \int_{x=M}^{\infty} \limsup_{t\to\infty} \mathbb{P}(Q(t) > x)dx
\end{aligned} \tag{120}$$

Combining Lemma 20 with (120), the proof is complete.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell system technical journal*, vol. 54, no. 8, pp. 1334–1387, 1975.

[2] S. Leung-Yan-Cheong and M. Hellman, "The gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[3] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, 2008.

[4] Y. Abdallah, M. A. Latif, M. Youssef, A. Sultan, and H. El Gamal, "Keys through arq: Theory and practice," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 737–751, 2011.

[5] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[6] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[7] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[8] K. Khalil, O. O. Koyluoglu, H. E. Gamal, and M. Youssef, "Opportunistic secrecy with a strict delay constraint," *arXiv preprint arXiv:0907.3341*, 2009.

[9] C. E. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[10] J. Barros and M. Bloch, "Strong secrecy for wireless channels (invited talk)," in *Information Theoretic Security*, pp. 40–53, Springer, 2008.

[11] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.

[12] A. C. Nascimento, J. Barros, S. Skludarek, and H. Imai, "The commitment capacity of the gaussian channel is infinite," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2785–2789, 2008.

[13] T. M. Cover and J. A. Thomas, *Elements of information theory*. 2nd ed. Wiley-interscience, 2006.

[14] J. Luo, L. Lin, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 323–330, 2003.

[15] S. Asmussen, S. Asmussen, and S. Asmussen, *Applied probability and queues*, vol. 2. Springer New York, 2003.

[16] P. Bertail and S. Clémençon, "A renewal approach to markovian u-statistics," *Mathematical Methods of Statistics*, vol. 20, no. 2, pp. 79–105, 2011.

[17] J. Kingman, "On queues in heavy traffic," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 383–392, 1962.

[18] R. Loynes, "The stability of a queue with non-independent inter-arrival and service times," in *Proc. Cambridge Philos. Soc*, vol. 58, pp. 497–520, Cambridge Univ Press, 1962.

[19] H. S. Kim and N. B. Shroff, "On the asymptotic relationship between the overflow probability and the loss ratio," *Advances in Applied Probability*, vol. 33, no. 4, pp. 836–863, 2001.

Onur Gungor received the B.S. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 2008, and the M.S. degree in electrical and computer engineering from Ohio State University, Columbus, in 2011. He interned with AT&T Labs, San Ramon, CA during spring and summer of 2012. He is currently a Ph.D. candidate at Ohio State University. Mr. Gungor is a recipient of Ohio State University Fellowship Award (2008).

Jian Tan received the B.S. degree from University of Science and Technology of China, Hefei, China in 2002, and the M.S. and Ph.D. degrees from Columbia University, New York, NY, USA in 2004 and 2009, respectively, all in electrical engineering. He is a Research Staff Member with IBM T. J. Watson Research Center, Yorktown Heights, NY, USA from 2010 to now. He was a postdoctoral researcher with the Networking and Communications Research Lab, The Ohio State University, Columbus, OH, USA, from 2009 to 2010. He interned with Lucent Bell Laboratories, Murray Hill, NJ, during the summers of 2005 and 2006, and with Microsoft Research, Cambridge, UK, in the winter of 2007. His current research interests focus on big data platforms, distributed computing, and related applications.

Dr. Tan was awarded the Eliahu Jury Award for his Ph.D. thesis work from Columbia University. He received the Best Student Paper Award at the 20th International Teletraffic Congress and the Best Paper Award at the 3rd IEEE International conference on Distributed Computing in Sensor Systems.

C. Emre Koksal (M‘03-SM‘12) received the B.S. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1996, and the S.M. and Ph.D. degrees in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge, in 1998 and 2002, respectively. He was a Postdoctoral Researcher in the Electrical Engineering and Computer Science Department, MIT, and in the School of Communication and Computer Sciences, EPFL, Lausanne, Switzerland, until 2006. Since then, he has been an Assistant Professor in the Electrical and Computer Engineering Department at The Ohio State University. His general areas of interest are wireless communication, information theory, and communication networks.

Dr. Koksal is the recipient of the National Science Foundation CAREER Award, the OSU College of Engineering Lumley Research Award, and the co-recipient of an HP LabsInnovation Research Award, all in 2011. The paper he coauthored was a best student paper candidate in ACM MobiCom 2005. Since 2013, he has been an Assciate Editor for IEEE Transactions on Wireless Communications and Elsevier Computer Networks.

Hesham El Gamal (M‘99-SM‘03-F‘10) received the B.S. and M.S. degrees in electrical engineering from Cairo University, Cairo, Egypt, in 1993 and 1996, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Maryland at College Park, MD, in 1999. From 1993 to 1996, he served as a Project Manager in the Middle East Regional Office of Alcatel Telecom. From 1996 to 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, the University of Maryland at College Park, MD. From February 1999 to December 2000, he was with the Advanced Development Group, Hughes Network Systems (HNS), Germantown, MD, as a Senior Member of Technical Staff. Since January 2001, he has been with the Electrical and Computer Engineering Department at Ohio State University where he is now a Professor. He held visiting appointments at UCLA, Institut Eurecom, and served as a Founding Director for the Wireless Intelligent Networks Center (WINC) at Nile University (2007-2009).

Dr. El Gamal is a recipient of the HNS Annual Achievement Award (2000), the OSU College of Engineering Lumley Research Award (2003, 2008), the OSU Electrical Engineering Department FARMER Young Faculty Development Fund (2003-2008), the OSU Stanley E. Harrison Award (2008), and the National Science Foundation CAREER Award (2004). He holds 12 patents and has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS (2001-2005), an Associate Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING (2003-2007), a Guest Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY SPECIAL ISSUE ON COOPERATIVE COMMUNICATIONS (2007), a member of the SP4COM technical committee (2002-2005), a cochair of the Globecom08 Communication Theory Symposium, and a co-chair of the 2010 IEEE Information Theory Workshop. He served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and a Guest Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Ness B. Shroff (S‘91-M‘93-SM‘01-F‘07) received his Ph.D. degree in Electrical Engineering from Columbia University in 1994. He joined Purdue university immediately thereafter as an Assistant Professor in the school of Electrical and Computer Engineering. At Purdue, he became Full Professor of ECE in 2003 and director of CWSA in 2004, a university-wide center on wireless systems and applications. In July 2007, he joined The Ohio State University, where he holds the Ohio Eminent Scholar endowed chair in Networking and Communications, in the departments of ECE and CSE. From 2009-2012, he served as a Guest Chaired professor of Wireless Communications at Tsinghua University, Beijing, China, and currently holds an honorary Guest professor at Shanghai Jiaotong University in China. His research interests span the areas of communication, social, and cyberphysical networks. He is especially interested in fundamental problems in the design, control, performance, pricing, and security of these networks. Dr. Shroff is a past editor for IEEE/ACM Trans. on Networking and the IEEE Communication Letters. He currently serves on the editorial board of the Computer Networks Journal, IEEE Network Magazine, and the Networking Science journal. He has chaired various conferences and workshops, and co-organized workshops for the NSF to chart the future of communication networks. Dr. Shroff is a Fellow of the IEEE and an NSF CAREER awardee. He has received numerous best paper awards for his research, e.g., at IEEE INFOCOM 2008, IEEE INFOCOM 2006, Journal of Communication and Networking 2005, Computer Networks 2003 (two of his papers also received runner-up awards at IEEE INFOCOM 2005 and INFOCOM 2013), and also student best paper awards (from all papers whose first author is a student) at IEEE WiOPT 2013, IEEE WiOPT 2012 and IEEE IWQoS 2006.