Ergodic Secret Message Capacity of the Wiretap Channel with Finite-Rate Feedback

Zouheir Rezki, *Senior Member, IEEE,* and Ashish Khisti, *Senior Member, IEEE,* and Mohamed-Slim Alouini, *Fellow, IEEE,*

***Abstract*—We study the secret message capacity of an ergodic block fading wiretap channel with partial channel state information at the transmitter and perfect channel state information at the receivers, under both a short term power constraint (STPC) and a long term power constraint (LTPC). We consider that in addition to the statistics of the main and the eavesdropper channel state information (CSI), the sender is provided by the legitimate receiver with a q-bit feedback, at the beginning of each coherence block, through an error-free public channel, with capacity q bits. We establish upper and lower bounds on the secrecy capacity. We show that the lower and the upper bounds coincide asymptotically as $q \to \infty$. When applied to Rayleigh fading channels, we show that, a 4-bit feedback achieves about 90% of the secrecy capacity when perfect main CSI is available at the transmitter. Finally, asymptotic analysis at high and low Signal-to-Noise Ratio (SNR) is presented. It is found that the capacity is bounded at high-SNR, whereas at asymptotically low-SNR, the lower bounds and the upper bound scale linearly with SNR under STPC. Furthermore, subject to LTPC, the capacity at low-SNR is equal to the capacity of the main channel without secrecy constraint and with perfect CSI at both the transmitter and the receiver, under a mild condition on the fading statistics. We also show that a positive secrecy rate is achievable even when the feedback is at the end of each coherence block and $q = 1$.**

***Index Terms*—Secrecy capacity, proactive feedback, ARQ feedback, high-SNR, low-SNR, capacity of fading channels.**

I. Introduction

The role of fading in providing physical layer security has been extensively highlighted recently, e.g. [1], [2]. In this context, we consider a wiretap channel consisting of a sender (Alice), a legitimate receiver (Bob) and an eavesdropper (Eve). Alice wants to communicate a secret message to Bob while keeping Eve in full ignorance of such a message. The main channel, between Alice and Bob, and the eavesdropper channel, between Alice and Eve, are both block fading channels. For arbitrarily large coherence blocks, and assuming we can code over sufficiently many coherence periods, and if the sender is perfectly aware of either the main CSI or both the main and the eavesdropper CSI, the secrecy capacity has been derived in [3]. Should the main CSI be imperfect at the sender, the secrecy capacity is still not known. In [4], using the so-called variational distance as a secrecy criterion, a single letter characterization of the secrecy capacity of an arbitrary wiretap channel with causal CSI at both the transmitter and the receivers is provided. However, this characterization is not easily computable due to the large space over which the prefixed random processes are taken, and thus the capacity expression therein turns out to be useful only to derive achievable rates. Independently and concurrently with our work [5], achievable rates for the ergodic and the block ergodic fading wiretap channel have also been derived in [6]. In [7], a lower and an upper bounds on the secrecy capacity are derived for a class of independent identically distributed (i.i.d.) fast fading channels, when the codeword length spans many coherence periods and when the sender has imperfect main CSI. Schemes based on sending an artificial noise to enhance the eavesdropper equivocation are presented in, e.g., [8]–[12]. Discussions on the effect of CSI estimation error on secrecy are also presented in [13]–[19].

In this paper, we assume that Bob knows its own channel instantaneously and Eve knows both its own channel and the main channel, instantaneously; whereas Alice is only aware of the statistics of these channels. There is also an error-free public feedback channel with limited capacity from Bob to Alice that may be tracked by Eve. In our setting, the feedback is exclusively used to send the output of a deterministic function that describes the main channel state information. The secret message capacity of this channel is not known. Several previous works have highlighted the impact of limited-rate feedback on the capacity of fading channels without secrecy constraint, see for instance, [20]–[23] and references therein. However, to the best of our knowledge, not much attention has been given to secret message capacity with limited-rate feedback.

For the setting described above, we first extend the scheme in [3] to incorporate q-bit feedback and observe numerically that when Rayleigh fading channels are considered, a 4-bit feedback achieves 90% of the secrecy capacity when perfect main CSI is available at the transmitter. As $q \to \infty$, we prove that this achievable rate coincides with the secrecy capacity when perfect main channel CSI is available at the transmitter. Then, we provide another achievable rate based on the wiretap code in [24], but accounting for the fading and the proactive feedback mechanism. We show that under a long term power constraint (LTPC), it achieves the capacity at asymptotically low SNR regime, with $q = 1$. Then, we establish a genie-

Zouheir Rezki and Mohamed-Slim Alouini are with the Electrical Engineering Program, Computer, Electrical, and Mathematical Science and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia. Email: {zouheir.rezki,slim.alouini}@kaust.edu.sa. Ashish Khisti is with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada, Email: {akhisti}@comm.utoronto.ca.

Part of this work has been presented at the 2012 IEEE International Symposium on Information Theory (ISIT'2012), Cambridge, MA, USA.

This publication was made possible by NPRP grant 5-603-2-243 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

aided upper bound similar to the one in e.g., [25], to include q-bit feedback. Since the transmitter is not aware of the main channel gain, the proof of the upper bound has required some extra technical steps following a similar approach as the one in [26], with proper adaptation to secrecy. Finally, we also show that even when the feedback is at the end of each coherence block and $q = 1$, a positive secrecy rate is achievable. This complements a result in [27] where a secret-key generation mechanism with 1-bit feedback at the end of the coherence block is proposed.

We then specialize our results to a short term power constraint (STPC) and a LTPC, respectively. We formulate the achievable rates and the upper bound as different optimization problems and propose an algorithm that attempts to find the optimal solution iteratively. In both cases, we present asymptotic analysis at high-SNR and low-SNR and show that in contrast to the high-SNR regime where the capacity is bounded; at asymptotically low-SNR regime, the lower bounds and the upper bound scale linearly with SNR under STPC. Since the upper bound is strictly smaller than the capacity without secrecy constraint, we argue that the secrecy induces a penalty even at low-SNR under STPC. On the contrary, under LTPC, the secrecy capacity is asymptotically equal to the capacity as if there is no secrecy constraint for a wide class of fading channels. Furthermore, we present a simple on-off scheme that is asymptotically (at low-SNR) capacity-achieving and that only requires 1-bit feedback in each coherence block.

The paper is organized as follows. Section II describes the system model and related background. Achievable rates and an upper bound on the secrecy capacity with finite rate feedback at the beginning and at the end of each coherence block are presented in Section III. In Section IV, we apply our results considering a STPC. In Section V, we analyze the ergodic secrecy capacity under LTPC, study the related optimization problems and propose an iterative algorithm to find the optimal solutions. In both Sections IV and V, asymptotic analysis at high and at low SNR are given under STPC and LTPC, respectively. Numerical results are reported in Section VI. Finally, Section VII concludes the paper.

Notations: The expectation operation is denoted by $\mathbf{E}[\cdot]$. The symbol $|x|$ is the modulus of the scalar x, while $[x]^+ = \max(0, x)$. The logarithms $\log(x)$ is the natural logarithm of x. We say that $f(x) \overset{a}{\approx} g(x)$ if and only if $\lim_{x\to a} \frac{f(x)}{g(x)} = 1$. When it is clear from the context, we omit a in $\overset{a}{\approx}$ for convenience.

II. System Model

We consider a discrete-time memoryless wire-tap channel consisting of a transmitter (Alice), a legitimate receiver (Bob) and an eavesdropper (Eve). Each terminal is equipped with a single antenna, i.e., a single-input single-output single-eavesdropper case. Alice wants to communicate a confidential messages W to Bob in the presence of Eve. The jth outputs at both the legitimate destination and the eavesdropper, at coherence period i, $i = 1, \ldots, L$, are expressed, respectively by:

$$\begin{cases} Y(i,j) = h(i)\, X(i,j) + U(i,j) \\ Z(i,j) = g(i)\, X(i,j) + V(i,j), \end{cases} \tag{1}$$

where $j = 1, \ldots, m$, with m representing the number of symbols in each coherence block; where $X(i,j) \in \mathbb{C}$ is the jth transmitted symbol at time coherence i, and $h(i) \in \mathbb{C}$, $g(i) \in \mathbb{C}$ are zero-mean and unit-variance channel gains that represent the main channel and the eavesdropper channel at time coherence i, respectively; and $U(i,j) \in \mathbb{C}$, $V(i,j) \in \mathbb{C}$ are zero-mean, unite-variance circularly symmetric white Gaussian noises. The fading process $\{h_i\}$ (resp. $\{g_i\}$) is assumed to be independent and identically distributed (i.i.d.) across the coherence blocks. Furthermore, the channel gains h and g are assumed to be independent of each other in any coherence interval. We assume perfect CSI at the receiver sides. That is, the legitimate receiver knows the instantaneous channel realizations $h(i)$, whereas the eavesdropper is aware of both $h(i)$ and $g(i)$. For convenience, we let $\gamma_h = |h|^2$, $\gamma_g = |g|^2$, $f_{\gamma_h}(\cdot)$ and $f_{\gamma_g}(\cdot)$ their PDF's, and $F_{\gamma_h}(\cdot)$, $F_{\gamma_g}(\cdot)$ their cumulative distribution functions (CDF).

The transmitter is not aware of neither $h(i)$ nor $g(i)$. However, in addition to the statistics of both channels, the transmitter is provided a q-bit (q integer and $q \geq 1$) feedback at the beginning (or at the end) of each coherence block, through an error-free public channel with limited capacity that is available to Alice and tracked by Eve. The feedback link is used to describe the main channel gain γ_h to the transmitter. More specifically, the channel gain support is split into 2^q intervals $[0, \tau_1), [\tau_1, \tau_2), \ldots, [\tau_{2^q-1}, \infty)$ and Bob sends back to Alice the output of a deterministic function $\kappa(\cdot)$ defined by: $\kappa(h) = k$, if $\gamma_h \in [\tau_k, \tau_{k+1})$. The result of this feedback drives Alice's decision to either transmit with the highest fixed rate R_k, $k = 1, \ldots, N$, where $N = 2^q - 1$, such that $R_k < \log(1 + \gamma_h(i)\, P)$ if such R_k exists, or to remain idle otherwise. Furthermore, the source is constrained to either a short term power constraint (STPC): $\mathbf{E}\left[\frac{1}{m}\sum_{j=1}^{m} |X(i,j)|^2\right] \leq P_{max}$, or to a long term power constraint (LTPC): $\mathbf{E}\left[\frac{1}{L}\sum_{i=1}^{L}\frac{1}{m}\sum_{j=1}^{m} |X(i,j)|^2\right] \leq P_{max}$.

We note that assuming statistical knowledge of the eavesdropper's channel may seem somehow not reasonable since the eavesdropper is passive and does not transmit. However, in case the eavesdropper belongs to the network, which is the case in this paper, the later assumption can be justified quite reasonably. For instance, one can think of a local area network (LAN) inside a secure building. The building is engineered such that artificial noise is injected isotropically in all directions. As a result of noise injection, the channel between each mobile (either legitimate or eavesdropper) and the access point is roughly the same on average. This would allow the transmitter to estimate the statistics of all nodes with high accuracy. Note that the idea of using artificial noise in order to create specific channel conditions that are favorable to secrecy has been already proposed in the literature, e.g., [28].

We are interested in message transmission secrecy capacity of such a channel when both L and m are sufficiently large. For convenience, let $n = mL$. The level of uncertainty about the message w at the eavesdropper is measured by the equivocation rate defined by:

$$R_e := \frac{1}{n} H\left(W \mid \boldsymbol{Z}^n, \boldsymbol{g}^L, \boldsymbol{h}^L\right) \tag{2}$$

where $H\left(W \mid \boldsymbol{Z}^n, \boldsymbol{g}^L, \boldsymbol{h}^L\right)$ denotes the conditional entropy of W given $\boldsymbol{Z}^n$, $\boldsymbol{g}^L$ and $\boldsymbol{h}^L$ and where $\boldsymbol{Z}^n = \left(Z(1,1), \ldots, Z(1,m), \ldots, Z(L,1), \ldots, Z(L,m)\right)$, $\boldsymbol{h}^L = \left(h(1), \ldots, h(L)\right)$ and $\boldsymbol{g}^L$ is defined similarly. The eavesdropper is ignorant about the message if $\lim_{n\to\infty} \frac{1}{n} I\left(W; \boldsymbol{Z}^n, \boldsymbol{g}^L, \boldsymbol{h}^L\right) = 0$, where $I(\cdot\,;\,\cdot)$ denotes the mutual information. A rate R is an achievable secrecy rate if for all $\epsilon > 0$, there exists a sequence of $\left(n, 2^{nR}, P_e\right)$ codes, for which 2^{nR} represents the number of messages to be sent to the destination, such that $R_e \geq R - \epsilon$ and $P_e \leq \epsilon$, where P_e is the average error probability defined by:

$$P_e = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \Pr\left\{W \neq \hat{W} \mid W = w\right\}, \tag{3}$$

where $\hat{W}$ is the output of the decoder at the intended receiver as a result of observing $\boldsymbol{Y}^n$. Furthermore, the secrecy capacity is given by: $C_s \coloneqq \sup_{R \in \mathcal{R}_s} R$, where $\mathcal{R}_s$ is the set of achievable secrecy rates.

III. Upper and Lower Bounds on the Secrecy Capacity

In this section, achievable rates and an upper bound are provided in Theorem 1 and Theorem 2, respectively. The proofs of these theorems are relegated to Appendix A and B, respectively.

Theorem 1: Let $\Pi^{(N)}$ be the set of all discrete power policies $\{P_k\}_{k=1}^N$ that satisfy the STPC (resp. LTPC). Let $\Theta^{(N)}$ be the set of all reconstruction points $\{\tau_k \mid 0 \leq \tau_1 \leq \ldots \leq \tau_N\}_{k=1}^N$ describing γ_h. For discrete-time memoryless channel described by (1), with an error-free q-bit feedback link at the beginning of each coherence block, the following rates are achievable:

$$R_{-1} = \max_{\substack{\{P_k\}_{k=1}^N \in \Pi^{(N)} \\ \{\tau_k\}_{k=1}^N \in \Theta^{(N)}}} \sum_{k=1}^{N} \Pr\left\{\tau_k \leq \gamma_h < \tau_{k+1}\right\} \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right)\right]^+\right] \tag{4}$$

$$R_{-2} = \max_{\substack{\{P_k\}_{k=1}^N \in \Pi^{(N)} \\ \{\tau_k\}_{k=1}^N \in \Theta^{(N)}}} \sum_{k=1}^{N} \Pr\left\{\tau_k \leq \gamma_h < \tau_{k+1}\right\} \cdot \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\log\left(\frac{1+\gamma_h P_k}{1+\gamma_g P_k}\right) \middle| \gamma_h \in [\tau_k, \tau_{k+1}]\right], \tag{5}$$

where for convenience, we set $\tau_{N+1} = \infty$.

Proof: The proof is given in Appendix A. ∎

It is particularly appealing to see that the lower bound in (4) is the sum over all possible rates of the product of the probability of success times the average rates gleaned by Bob over Eve during all fading realizations. A similar fact has been observed in [27], but for secret key sharing (not message transmission) and for $N = 1$. Theorem 1 states that a 1-bit feedback at the beginning of each coherence block guarantees a positive secrecy rate. We now present an upper bound on the secrecy rate with proactive feedback.

Theorem 2 (Upper bound): Let $\Pi_{(0)}^{(N)}$ be the set of all power policies $\{P_k\}_{k=0}^N$ that satisfy the STPC (resp. LTPC). Let $\Theta_{(0)}^{(N)}$ be the set of all reconstruction points $\{\tau_k \mid 0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_N\}_{k=0}^N$ describing γ_h. For the discrete-time memoryless channel described by (1), with an error-free q-bit feedback link at the beginning of each coherence block, an upper bound on the secrecy capacity is given by:

$$R_{+} = \max_{\substack{\{P_k\}_{k=0}^N \in \Pi_{(0)}^{(N)} \\ \{\tau_k\}_{k=0}^N \in \Theta_{(0)}^{(N)}}} \sum_{k=0}^{N} \Pr\left\{\tau_k \leq \gamma_h < \tau_{k+1}\right\} \cdot \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\left[\log\left(\frac{1+\gamma_h P_k}{1+\gamma_g P_k}\right)\right]^+ \middle| \gamma_h \in [\tau_k, \tau_{k+1}]\right], \tag{6}$$

where for convenience, we set $\tau_{N+1} = \infty$. Furthermore, R_{-1} in Theorem 1 coincides with R_+ as $N \to \infty$.

Proof: The proof is given in Appendix B. ∎

The lower bound in (4) is an increasing function of N and as shown in Appendix B, the lower bound R_{-1} in (4) and the upper bound R_+ in (6) match as $N \to \infty$, hence fully characterizing the ergodic capacity in this case. Letting N goes to ∞ may be interpreted as if there is a noiseless public link with infinite capacity for which the secrecy capacity is given by [3]:

$$R_{++} = \max_{P(h)} \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\left[\log\left(\frac{1+\gamma_h P(h)}{1+\gamma_g P(h)}\right)\right]^+\right]. \tag{7}$$

Letting N goes to ∞ may seem too restrictive as our feedback link is of limited-capacity. Fortunately, this asymptotic behavior starts showing up for relatively small N values as shown by our numerical results, i.e., it takes only few feedback bits to achieve most of the available secrecy capacity, at least in a Rayleigh fading scenario. It is worth mentioning that guaranteeing a positive secrecy rate is not really tied to knowing the feedback at the **beginning** of each coherence block. Providing a feedback at the **end** of each coherence block instead, also guarantees a positive secrecy rate, although smaller than the one given by (4). To see this, let us assume that at the end of each coherence block, Bob feeds back a 1-bit ARQ to Alice informing her whether the actual frame has been correctly decoded (ACK), or not (NACK). Alice keeps retransmitting the same block until she gets an ACK, then moves on to the next frame. Clearly, because some of the frames are transmitted more than once, this scheme leaks some information to the eavesdropper. Ultimately, one can assume that the blocks repeated because of the NACK feedback are completely revealed to the eavesdropper as a worst-case scenario. Fortunately, even such a conservative scheme guarantees a positive secrecy rate as formalized in Theorem 3.

Theorem 3: A lower bound on the secrecy capacity of the discrete-time memoryless channel described by (1), with an error-free 1-bit ARQ feedback at the end of each coherence block, is given by:

$$R_{--} = \max_{\substack{\{P\} \in \Pi^{(1)} \\ \{\tau\} \in \Theta^{(1)}}} \theta^2 \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{1+\tau P}{1+\gamma_g P}\right)\right]^+\right], \tag{8}$$

where θ is the probability of success defined by $\theta = \Pr\{\gamma_h \geq \tau\}$. The upper bound in (6), with $N = 1$, still holds.

Proof: The proof is presented in Appendix C. ∎

While the rate in Theorem 3 only accounts for the contribution of the blocks that have not been repeated into the secrecy rate, it can be immediately improved by accounting for the contribution of the blocks that have been repeated, say once, into the secrecy rate.

Corollary 1: A lower bound on the secrecy capacity of the discrete-time memoryless channel described by (1), with an error-free 1-bit ARQ feedback at the end of each coherence block, is given by:

$$\begin{aligned} R_{--}^{+} \;=\; & \max_{\substack{\{P\}\in\Pi^{(1)}\\ \{\tau\}\in\Theta^{(1)}}} \left\{ \theta^2 \underset{\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau P}{1+\gamma_g P} \right) \right]^{+} \right] \right. \\ & \left. + \theta^2 \;(1-\theta) \underset{\gamma_g^{(2)}}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau P}{1+\gamma_g^{(2)} P} \right) \right]^{+} \right] \right\}, \end{aligned} \tag{9}$$

where $\gamma_g^{(2)}$ is a random variable distributed as the sum of two independent γ_g's.

Proof: The proof is presented in Appendix D ∎

We now apply our results considering STPC and LTPC, respectively.

IV. Ergodic Capacity under STPC

We note that under STPC, (4), (5), (6) and (8) induce different optimization problems which we designate as $\mathcal{P}_i$, $i = 1,\ldots,4$, for convenience. For $\mathcal{P}_i$'s, $i \in \{1,3,4\}$, it is easy to see that the optimal power policy consists of setting all power equal to P_{max}. This follows from the fact that $x \mapsto \left[\log\left(\frac{1+a\,x}{1+b\,x}\right)\right]^{+}$ is a non-decreasing function for all $x \geq 0$ and all a and b reals. Hence, the upper bounds given by (6) and (7) are equal under STPC and we have:

$$R_{+} = R_{++} = \underset{\gamma_h,\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\gamma_h P_{max}}{1+\gamma_g P_{max}} \right) \right]^{+} \right]. \tag{10}$$

Consequently, although our upper bound R_{++} has been established by providing q-bit feedback only to the transmitter, it does not improve over the secrecy capacity under perfect main CSI due to STPC. Moreover, for $\mathcal{P}_1$ and $\mathcal{P}_4$, the optimal reconstruction points $\left\{\tau_k^*\right\}$ are obtained by solving the Karush Kuhn Tucker (KKT) conditions which are necessary conditions only due to the non-convexity of $\mathcal{P}_1$ and $\mathcal{P}_4$ in τ_k's.[1]. Below, we show calculation details of the KKT condition related to $\mathcal{P}_1$. We first form the Lagrangian as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\tau},\boldsymbol{\lambda}) = & \sum_{k=1}^{N} \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k) \right) \underset{\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k} \right) \right]^{+} \right] \\ & - \sum_{k=0}^{N-1} \lambda_k (\tau_k - \tau_{k+1}) \end{aligned} \tag{11}$$

where $\boldsymbol{\tau} = (\tau_0,\ldots,\tau_{N+1})$, with $\tau_0 = 0$ and $\tau_{N+1} = \infty$ and where $\boldsymbol{\lambda} = (\lambda_0,\ldots,\lambda_{N-1})$ is the vector of non-negative Lagrange multipliers corresponding to the constraints $\tau_k \leq \tau_{k+1}$, $k = 0,\ldots,N-1$. The KKT conditions imply that the partial derivative of $\mathcal{L}(\boldsymbol{\tau},\boldsymbol{\lambda})$ with respect to τ_k is equal to zero which yields:

$$\begin{aligned} & f_{\gamma_h}(\tau_k) \int_0^{\tau_{k-1}} \log\left(\frac{1+\tau_{k-1}P_{max}}{1+\gamma_g P_{max}} \right) f_{\gamma_g}(\gamma_g)\,\mathrm{d}\gamma_g \\ & - f_{\gamma_h}(\tau_k) \int_0^{\tau_k} \log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}} \right) f_{\gamma_g}(\gamma_g)\,\mathrm{d}\gamma_g + (\lambda_{k-1} - \lambda_k) \\ & + \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k) \right) \int_0^{\tau_k} \frac{P_{max}}{1+\tau_k P_{max}} f_{\gamma_g}(\gamma_g)\,\mathrm{d}\gamma_g \\ & = 0. \end{aligned} \tag{12}$$

We note that there is no loss of optimality by taking $0 < \tau_1 \ldots < \tau_N$ since if $\tau_k = \tau_{k+1}$ for some $k = 1,\ldots,N$, then the kth element of the sum in (4) is equal to zero and hence contributes nothing to the objective function. Since all τ_k's are different, then by the complimentary slackness conditions $\lambda_k(\tau_k - \tau_{k+1}) = 0$, all λ_k's are equal to zero. Using the later fact, the condition (12) simplifies to

$$\begin{aligned} & f_{\gamma_h}(\tau_k) \left(\underset{\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau_{k-1}P_{max}}{1+\gamma_g P_{max}} \right) \right]^{+} \right] - \underset{\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}} \right) \right]^{+} \right] \right) \\ & + \frac{\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k) \right) F_{\gamma_g}(\tau_k) P_{max}}{1+\tau_k P_{max}} \\ & = 0, \end{aligned} \tag{13}$$

for $k = 1,\ldots,N$. Along similar lines, the KKT condition for $\mathcal{P}_2$ can also be obtained as

$$2\, f_{\gamma_h}(\tau)\, \underset{\gamma_g}{\mathbf{E}} \left[\left[\log\left(\frac{1+\tau P_{max}}{1+\gamma_g P_{max}} \right) \right]^{+} \right] - \frac{\left(1 - F_{\gamma_h}(\tau)\right) F_{\gamma_g}(\tau) P_{max}}{1+\tau P_{max}} = 0. \tag{14}$$

For $\mathcal{P}_2$, the integrand in R_{-2} is not necessarily increasing in P_k and hence it is generally not clear whether P_{max} is optimal or not. In fact, it is not even clear whether the rate R_{-2} is non-negative for given P_{max} and N values. Nevertheless, for a class of fading for which given the CSI feedback, the main channel distribution becomes "more informative" than the eavesdropper's, the target in (5) is increasing in P_k and thus setting $P_k = P_{max}$, $k = 1,\ldots,N$, is optimal. The term "more informative" can be made more precise through stochastic dominance theory. Before, formalizing our result above, we need the following definition [29].

Definition 1: A random variable (r.v.) X is first-order stochastically dominant (FOSD) than a r.v. Y, denoted as $X \geq Y$, if $\Pr\{X \geq c\} \geq \Pr\{Y \geq c\}$, for every real c.

We now give a sufficient condition on the fading statistics under which the rate R_{-2} in (5) is non-negative and $P_k = P_{max}$, $k = 1,\ldots,N$, is optimal.

Lemma 1: Let a be an arbitrary real such that $a > 0$. Let $\gamma_{h,[a,\infty)}$ be the r.v. γ_h conditioned on the event $\gamma_h \in [a,\infty)$, i.e., $\gamma_{h,[a,\infty)} = \gamma_h \mid \gamma_h \in [a,\infty)$. Then, if $\gamma_{h,[a,\infty)} \geq \gamma_g$ for all $a > 0$, the following statements hold true for $\mathcal{P}_2$:

i) $P_k = P_{max}$, $k = 1,\ldots,N$, is optimal.
ii) R_{-2} in (5) is equal to:

$$R_{-2} = \Pr\{\gamma_h \geq \tau^*\} \underset{\gamma_h,\gamma_g}{\mathbf{E}} \left[\log\left(\frac{1+\gamma_h P_{max}}{1+\gamma_g P_{max}} \right) \middle| \gamma_h \in [\tau^*,\infty) \right], \tag{15}$$

[1]Note that under STPC, the optimizations considered have all affine constraints which is sufficient in order for the KKT conditions to provide necessary conditions.

where τ^* is given by:

$$\tau^* = \left(\exp\left(\mathop{\mathbf{E}}_{\gamma_g}\left[\log\left(1+\gamma_g P_{max}\right)\right]\right)-1\right)/P_{max}.$$

iii) $R_{-2} \geq 0$.

Proof: For convenience, the proof is presented in Appendix E. ∎

Lemma 1 states that for the class of fading where the CSI feedback renders the main channel FOSD than the eavesdropper's channel, R_{-2} is non-negative for any P_{max}. However, and as asserted by Lemma 1, the secrecy rate R_{-2} does not increase with N and thus providing more than 1 bit feedback to the source is unfortunately useless under STPC.

***Remark** 1:* If the main channel and the eavesdropper's channel are identically distributed, then for any $c \geq 0$ and any $a > 0$, we have:

$$\begin{aligned}
\Pr\{\gamma_{h,[a,\infty)} \geq c\} &= \int_c^\infty f_{\gamma_{h,[a,\infty)}}(x)\,\mathrm{d}x && (16)\\
&= \int_c^\infty \frac{f_{\gamma_h}(x)\,\mathbb{1}_{[a,\infty)}(x)}{1-F_{\gamma_h}(a)}\mathrm{d}x && (17)\\
&\geq \int_c^\infty f_{\gamma_h}(x)\,\mathrm{d}x && (18)\\
&= \Pr\left\{\gamma_g \geq c\right\} && (19)
\end{aligned}$$

where $\mathbb{1}_{[a,\infty)}(x)$ is an indicator function that is equal to 1 if $x \in [a,\infty)$ and 0 otherwise. and where (18) can be verified easily. That is, if γ_h and γ_g are identically distributed, then $\gamma_{h,[a,\infty)}$ is FOSD than γ_g. This implies that assuming that the fading are identically distributed is stronger than the assumption $\gamma_{h,[a,\infty)} \geq \gamma_g$ in Lemma 1.

In the sequel, we focus on fading γ_h and γ_g that satisfy the condition in Lemma 1, i.e., $\gamma_{h,[a,\infty)} \geq \gamma_g$, for all $a > 0$.

A. Asymptotic Analysis at High-SNR and Low-SNR

In this section, we assume that there is a q-bit feedback at the beginning of each coherence, i.e., the settings of Theorem 1 and we are interested in the secrecy capacity at asymptotically high-SNR and low-SNR regimes.

1) High-SNR Regime: Our result is summarized in Corollary 2.

Corollary 2: At high-SNR ($P_{max} \to \infty$), the secrecy capacity is bounded, i.e., does not grow with P_{max}. Furthermore, the following rates are achievable:

$$R_{-1}^{\infty} = \max_{0\leq\tau_1\leq\ldots\leq\tau_N} \sum_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right] \quad (20)$$

$$R_{-2}^{\infty} = \max_{\tau\geq 0} \mathop{\mathbf{E}}_{\frac{\gamma_h}{\gamma_g}\geq\tau}\left[\log\left(\frac{\gamma_h}{\gamma_g}\right)\right]. \quad (21)$$

An upper bound on the secrecy capacity is given by:

$$R_{+}^{\infty} = \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\left[\log\left(\frac{\gamma_h}{\gamma_g}\right)\right]^+\right]. \quad (22)$$

Proof: The proof is presented in Appendix G. ∎

To determine τ_k's in (20), we solve the necessary KKT conditions:

$$\begin{aligned}
&f_{\gamma_h}(\tau_k)\left(\mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_{k-1}}{\gamma_g}\right)\right]^+\right] - \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right]\right)\\
&+\frac{\left(F_{\gamma_h}(\tau_{k+1})-F_{\gamma_h}(\tau_k)\right)F_{\gamma_g}(\tau_k)}{\tau_k} = 0, \quad (23)
\end{aligned}$$

whereas the optimal τ^* in (21) is equal to:

$$\tau^* = \exp\left(\mathop{\mathbf{E}}_{\gamma_g}\left[\log\left(\gamma_g\right)\right]\right). \quad (24)$$

2) Low-SNR Regime: Motivated by the boundedness of the secrecy capacity at high-SNR, we analyze in this section the secrecy capacity at low-SNR regime. Our result is rather positive as it states that under STPC, the capacity is asymptotically (at low-SNR) linear in SNR as formalized in Corollary 3.

Corollary 3: At low-SNR ($P_{max} \to 0$), the secrecy capacity is linear in P_{max}. Furthermore, the following rates are achievable:

$$R_{-1} \approx P_{max} \cdot \max_{0\leq\tau_1\leq\ldots\leq\tau_N} \sum_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k-\gamma_g\right]^+\right] \quad (25)$$

$$R_{-2} \approx P_{max} \cdot \mathop{\mathbf{E}}_{\gamma_h\geq\mathbf{E}_{\gamma_g}[\gamma_g]}\left[\gamma_h - \mathop{\mathbf{E}}_{\gamma_g}[\gamma_g]\right]. \quad (26)$$

An upper bound on the secrecy capacity is given by:

$$R_{+} \approx P_{max} \cdot \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\left[\gamma_h-\gamma_g\right]^+\right]. \quad (27)$$

Proof: The proof is presented in Appendix H. ∎

Here again, to determine τ_k's in (25), we solve the necessary KKT conditions:

$$\begin{aligned}
&f_{\gamma_h}(\tau_k)\left(\mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_{k-1}-\gamma_g\right]^+\right] - \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k-\gamma_g\right]^+\right]\right)\\
&+\left(F_{\gamma_h}(\tau_{k+1})-F_{\gamma_h}(\tau_k)\right)F_{\gamma_g}(\tau_k) = 0. \quad (28)
\end{aligned}$$

V. Ergodic Capacity under LTPC

Under LTPC, the results in Theorem 1, Theorem 2 and Theorem 3 remain valid, with the difference that the LTPC must be satisfied. More specifically, we identify the related optimization problems as follows:

$$\bar{\mathcal{P}}_1 : \begin{cases} \max\limits_{0\leq\tau_1\leq\ldots\leq\tau_N} \sum\limits_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} \cdot \mathop{\mathbf{E}}\limits_{\gamma_g}\left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right)\right]^+\right] \\ \text{s.t.} \sum\limits_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} P_k \leq P_{max}, \end{cases} \quad (29)$$

$$\bar{\mathcal{P}}_2 : \begin{cases} \max\limits_{0\leq\tau_1\leq\ldots\leq\tau_N} \sum\limits_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} \\ \cdot \mathop{\mathbf{E}}\limits_{\gamma_h,\gamma_g}\left[\log\left(\frac{1+\gamma_h P_k}{1+\gamma_g P_k}\right)\middle|\, \gamma_h \in [\tau_k,\tau_{k+1}]\right] \\ \text{s.t.} \sum\limits_{k=1}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} P_k \leq P_{max}, \end{cases} \quad (30)$$

$$\bar{\mathcal{P}}_3 : \begin{cases} \max\limits_{0\leq\tau_1\leq\ldots\leq\tau_N} \sum\limits_{k=0}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} \\ \cdot \mathop{\mathbf{E}}\limits_{\gamma_h,\gamma_g}\left[\left[\log\left(\frac{1+\gamma_h P_k}{1+\gamma_g P_k}\right)\right]^+\middle|\, \gamma_h \in [\tau_k,\tau_{k+1}]\right] \\ \text{s.t.} \sum\limits_{k=0}^{N} \Pr\{\tau_k \leq \gamma_h < \tau_{k+1}\} P_k \leq P_{max} \end{cases} \quad (31)$$

$$\bar{\mathcal{P}}_4 : \begin{cases} \max\limits_{\tau \geq 0} \Pr\{\gamma_h \geq \tau\}^2 \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{1+\tau P}{1+\gamma_g P}\right)\right]^+\right]. \\ \text{s.t. } \Pr\{\gamma_h \geq \tau\} P \leq P_{max}, \end{cases} \tag{32}$$

where the acronym s.t. stands for "subject to". Although $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_3$ and $\bar{\mathcal{P}}_4$ are convex in P_k's, none of the above optimization problems is convex in τ_k's and hence they are all non-convex. However, one can focus again on the dual problem and relies again on the KKT conditions that provide necessary conditions assuming a ceratin qualification constraint at the maximizers [30]. Similarly to STPC case, there is no loss of optimality by taking $0 < \tau_1 \ldots < \tau_N$ since if $\tau_k = \tau_{k+1}$ for some $k = 1, \ldots, N$, then the kth element in the sum of the objective function in $\bar{\mathcal{P}}_1$ is equal to zero and hence contributes nothing to the objective function. For convenience, the related KKT conditions for $\bar{\mathcal{P}}_i$, $i = 1, \ldots, 4$, are provided in Table I, where $\frac{\partial}{\partial P_k}$ and $\frac{\partial}{\partial \tau_k}$ represent the derivative of the dual objective function with respect to P_k's and τ_k's, respectively; and where μ is the Lagrange multiplier associated with LTPC. Below, we show calculation details of the KKT conditions for $\bar{\mathcal{P}}_1$, similar derivations are used for other $\bar{\mathcal{P}}_i$'s, $i = 2, 3, 4$. We first form the Lagrangian as

$$\begin{aligned} \mathcal{L}(\boldsymbol{P}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mu) = & \sum_{k=1}^{N} \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right)\right]^+\right] \\ & -\mu\left(\sum_{k=1}^{N}\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) P_k - P_{max}\right) \\ & - \sum_{k=0}^{N-1} \lambda_k (\tau_k - \tau_{k+1}) \end{aligned} \tag{33}$$

where $\boldsymbol{P} = (P_1, \ldots, P_N)$; where $\boldsymbol{\tau} = (\tau_0, \ldots, \tau_{N+1})$, with $\tau_0 = 0$ and $\tau_{N+1} = \infty$ and where $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_{N-1})$ is the vector of non-negative Lagrange multipliers corresponding to the constraints $\tau_k \leq \tau_{k+1}$, $k = 0, \ldots, N-1$. The KKT conditions imply that the partial derivative of $\mathcal{L}(\boldsymbol{P}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mu)$ with respect to P_k is equal to zero which yields:

$$\begin{aligned} & \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) \int_0^{\tau_k} \left(\frac{\tau_k}{1+\tau_k P_k} - \frac{\gamma_g}{1+\gamma_g P_k}\right) f_{\gamma_g}(\gamma_g)\, \mathrm{d}\gamma_g \\ & -\mu\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) \\ & = 0. \end{aligned} \tag{34}$$

Again, there is no loss of optimality in considering that $F_{\gamma_h}(\tau_1) < \ldots < F_{\gamma_h}(\tau_N)$ since if $F_{\gamma_h}(\tau_k) = F_{\gamma_h}(\tau_{k+1})$ for some $k = 1, \ldots, N$, then the kth element in the sum of the objective function in $\bar{\mathcal{P}}_1$ contributes nothing to the objective function. Hence, simplifying (34) yields the $\frac{\partial}{\partial P_k}$ condition in Table I. Similarly, taking the partial derivative of $\mathcal{L}(\boldsymbol{P}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mu)$ with respect to τ_k yields:

$$\begin{aligned} & f_{\gamma_h}(\tau_k) \int_0^{\tau_{k-1}} \log\left(\frac{1+\tau_{k-1}P_{k-1}}{1+\gamma_g P_{k-1}}\right) f_{\gamma_g}(\gamma_g)\, \mathrm{d}\gamma_g \\ & - f_{\gamma_h}(\tau_k) \int_0^{\tau_k} \log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right) f_{\gamma_g}(\gamma_g)\, \mathrm{d}\gamma_g \\ & + \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) \int_0^{\tau_k} \frac{P_k}{1+\tau_k P_k} f_{\gamma_g}(\gamma_g)\, \mathrm{d}\gamma_g \\ & -\mu f_{\gamma_h}(\tau_k)(P_{k-1} - P_k) - (\lambda_{k-1} - \lambda_k) \\ & = 0. \end{aligned} \tag{35}$$

Again, since all τ_k's are different, then by the complimentary slackness conditions $\lambda_k(\tau_k - \tau_{k+1}) = 0$, all λ_k's are equal to zero. Using the later fact, the condition $\frac{\partial}{\partial \tau_k}$ in Table I follows immediately from (35).

Analyzing closely the derivatives with respect to P_k's in Table I, it can be shown that $\mu > 0$ and thus the power constraint is satisfied with equality, for $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_3$ and $\bar{\mathcal{P}}_4$. Indeed, since $f_{\gamma_h}(\cdot)$ and $f_{\gamma_g}(\cdot)$ are continuous and necessarily positive in an interval inside $[\tau_k, \tau_{k+1}]$, for some $k = 1, \ldots, N$, (otherwise the objective function would be equal to zero), and since the arguments of the expectation function in the $\frac{\partial}{\partial P_k}$ condition for $\bar{\mathcal{P}}_1$, $\bar{\mathcal{P}}_3$ and $\bar{\mathcal{P}}_4$ are positive, then μ is necessarily positive. For $\bar{\mathcal{P}}_2$, it is not clear whether μ is positive or equal to zero and this seems to depend on the fading's PDF. Nevertheless, one can show that when the main and the eavesdropper channels are identically distributed, μ is in fact strictly positive.[2]

Solving the KKT conditions in a closed form is very challenging. Instead, we present below an iterative algorithm that attempts to find the optimal solution using the KKT conditions. A similar algorithm has been proposed in [21], but without secrecy constraint. Likewise in [21], we do not claim the convergence of Algorithm 1.

Algorithm 1 Secrecy Rate with Feedback under a Long Term Power Constraint (LTPC)

Initialize $i = 0$, $P_k^{(0)} = P_{max}$, $\forall\, i$, set $\mu^{(0)}$ arbitrarily;
repeat
  Fix $\left\{P_k^{(i)}\right\}$ and $\mu^{(i)}$, solve for $\left\{\tau_k^{(i)}\right\}$ using $\frac{\partial}{\partial \tau_k}$ in Table I;
  Compute R^i;
  Fix $\left\{\tau_k^{(i)}\right\}$, find $\left\{P_k^{(i+1)}\right\}$ and $\mu^{(i+1)}$ using $\frac{\partial}{\partial P_k}$ in Table I;
  $i \leftarrow i + 1$;
until Convergence: $\frac{R^{(i+1)} - R^{(i)}}{R^{(i+1)}} \leq \epsilon$;

In Algorithm 1, $R^{(i)}$ represents either R_{-1}, R_{-2}, R_+ or R_{--}, at the ith iteration; and ϵ is an arbitrary small positive number. Note that for $\bar{\mathcal{P}}_i$, $i = 1, 2, 3$, solving the $\frac{\partial}{\partial \tau_k}$ condition in Table I can be done recursively starting from $k = 1$ until $k = N$, using standard root finding algorithms.

A. Asymptotic Analysis at High-SNR and Low-SNR

Here again, we assume that there is a q-bit feedback at the beginning of each coherence, and we are interested in the secrecy capacity at asymptotically high-SNR and low-SNR regimes. While at high-SNR, the results in Corollary 2 still hold confirming that likewise without secrecy constraint, power adaptation does not provide any additional capacity gain at high-SNR under secrecy constraint; we show that at low-SNR, power adaptation drastically increases the achievable secrecy rate. More interestingly, we show that under LTPC, the capacity under secrecy constraint is asymptotically equal to the capacity as if there is no secrecy constraint, for fading channels with unbounded support. Moreover, 1-bit feedback

[2]The proof of this result is provided in Appendix F.

Table I
SUMMARY OF $\bar{\mathcal{P}}_i$'S KKT CONDITIONS, $i = 1, \ldots, 4$.

	$\frac{\partial}{\partial P_k}$	$\frac{\partial}{\partial \tau_k}$
$\bar{\mathcal{P}}_1$	$\underset{\gamma_g}{\mathbf{E}}\left[\left[\frac{\tau_k}{1+P_k\tau_k} - \frac{\gamma_g}{1+P_k\gamma_g}\right]^+\right] - \mu = 0$	$f_{\gamma_h}(\tau_k)\left(\underset{\gamma_g}{\mathbf{E}}\left[\left[\log\left(\frac{1+\tau_{k-1}P_{k-1}}{1+\gamma_g P_{k-1}}\right)\right]^+\right] - \underset{\gamma_g}{\mathbf{E}}\left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right)\right]^+\right] - \mu(P_{k-1}-P_k)\right) + \frac{\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right)F_{\gamma_g}(\tau_k)P_k}{1+\tau_k P_k} = 0$
$\bar{\mathcal{P}}_2$	$\underset{\tau_k\le\gamma_h\le\tau_{k+1}}{\mathbf{E}}\left[\frac{\gamma_h}{1+P_k\gamma_h}\right] - \left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right)\left(\underset{\gamma_g}{\mathbf{E}}\left[\frac{\gamma_g}{1+P_k\gamma_g}\right] + \mu\right) = 0$	$f_{\gamma_h}(\tau_k)\left(\log\left(\frac{1+P_{k-1}\tau_k}{1+P_k\tau_k}\right) - \underset{\gamma_g}{\mathbf{E}}\left[\log\left(\frac{1+\gamma_g P_{k-1}}{1+\gamma_g P_k}\right)\right] + \mu(P_k - P_{k-1})\right) = 0$
$\bar{\mathcal{P}}_3$	$\underset{\substack{\tau_k\le\gamma_h<\tau_{k+1}\\ \gamma_g}}{\mathbf{E}}\left[\left[\frac{\gamma_h}{1+\gamma_h P_k} - \frac{\gamma_g}{1+\gamma_g P_k}\right]^+\right] - \mu\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) = 0$	$f_{\gamma_h}(\tau_k)\left(\underset{\gamma_g}{\mathbf{E}}\left[\left[\log\left(\frac{1+\tau_k P_{k-1}}{1+\gamma_g P_{k-1}}\right)\right]^+\right] - \underset{\gamma_g}{\mathbf{E}}\left[\left[\log\left(\frac{1+\tau_k P_k}{1+\gamma_g P_k}\right)\right]^+\right] - \mu(P_{k-1}-P_k)\right) = 0$
$\bar{\mathcal{P}}_4$	$\left(1 - F_{\gamma_h}(\tau)\right)\underset{\gamma_g}{\mathbf{E}}\left[\left[\frac{\tau}{1+P\tau} - \frac{\gamma_g}{1+P\gamma_g}\right]^+\right] - \mu = 0$	$f_{\gamma_h}(\tau)\left(-2\left(1 - F_{\gamma_h}(\tau)\right)\underset{\gamma_g}{\mathbf{E}}\left[\left[\log\left(\frac{1+\tau P}{1+\gamma_g P}\right)\right]^+\right] + \mu P\right) + \frac{\left(1-F_{\gamma_h}(\tau)\right)^2 F_{\gamma_g}(\tau)P}{1+\tau P} = 0$

is enough to achieve this capacity. These statements are made precise in Theorem 4.

1) Low-SNR Regime:

Theorem 4: For fading channels with infinite support, the secrecy capacity at low-SNR, $C_s(P_{max})$, of the channel described by (1), with an error-free q-bit feedback link at the beginning of each coherence block is given by:

$$C_s(P_{max}) \overset{0}{\approx} C_{w.s}(P_{max}), \tag{36}$$

where $C_{w.s}(\cdot)$ stands for the capacity of the main channel without secrecy constraint and with perfect CSI at both the transmitter (CSI-T) and the receiver (CSI-R). Furthermore, 1-bit feedback at the beginning of each coherence block is enough to achieve this capacity.

Proof: The proof is presented in Appendix I. ■

Few remarks are worthwhile:

***Remark** 2:* The fact that the secrecy capacity is asymptotically equal to the capacity as if there is no secrecy constraint, further stresses on the value of CSI at the transmitter at low-SNR regime. Recall that with neither a feedback nor main CSI at the transmitter, the secrecy capacity is equal to zero. Theorem 4 highlights the fact that even with 1-bit feedback, not only one can achieve secrecy at low-SNR, but this secrecy is obtained with a vanishing capacity-penalty to the legitimate receiver due to the presence of the eavesdropper. Nevertheless, since our achievable rate R_{-2} follows from Csiszár-Körner characterization of the secrecy capacity [24], a wiretap code is still needed to bin the secret message.

***Remark** 3:* The encoding scheme related to R_{-2} exploits the advantage that the legitimate receiver has over the eavesdropper through the feedback link. As shown in [7], this scheme ensures (by properly optimizing over τ) a positive secrecy rate for an arbitrary P_{max} value. This hinges on the fact that if the main channel is "good", it is more unlikely that the eavesdropper's channel be better. The later heuristic statement can be proven rigorously by computing the probability that γ_g be better than γ_h, given that $\gamma_h \ge \tau$, as follows:

$$\begin{aligned}\Pr\left\{\gamma_g \ge \gamma_h \mid \gamma_h \ge \tau\right\} &= \frac{\Pr\left\{\gamma_g \ge \gamma_h,\, \gamma_h \ge \tau\right\}}{\Pr\{\gamma_h \ge \tau\}} \\ &= \frac{1 - F_{\gamma_h}(\tau) - \frac{1}{2}\left(1 - F^2_{\gamma_h}(\tau)\right)}{1 - F_{\gamma_h}(\tau)} \\ &= \frac{1}{2}\Pr\{\gamma_h \ge \tau\}.\end{aligned}$$

While this scheme is not necessarily the best strategy at an arbitrary P_{max} value, it is particularly pleasing to see that it is in fact enough to achieve the secrecy capacity at asymptotically low-SNR.

***Remark** 4:* The result in Theorem 4 relies on the fact that the main channel fading has an infinite support. Should the main channel have a finite support, G would be finite and the limit in (122) is not equal to zero. In fact, if the fading support is bounded, the result in Theorem 4 does not hold anymore.[3]

VI. NUMERICAL RESULTS

In this section, numerical results are provided for Rayleigh fading channels such that $\underset{\gamma_h}{\mathbf{E}}[\gamma_h] = \underset{\gamma_g}{\mathbf{E}}\left[\gamma_g\right] = 1$. Figure 1 depicts the lower bounds and the upper bound in Theorem 1 and 2 in nats per channel use (npcu) versus P_{max} (designated here as SNR), for different q-bit feedback scenarios. Also shown in Fig. 1 are the achievable rates R_{--} in Theorem 3 and R^+_{--} in Corollary 1 with 1-bit ARQ feedback. Figure 1 confirms the positive secrecy rate even for 1-bit feedback (at the beginning or at the end of each coherence block). As the number of feedback bits increases (here $q = 4$), the lower bound gets closer to the upper bound confirming the statement in Theorem 1 and about 90% of the upper bound is achieved for all SNR values displayed in Fig. 1. The high-SNR characterizations R^∞_{-2} and R^∞_{+} given in Corollary 2 are also plotted in Fig. 1 where it can be seen that the relative gap between them is of order 30%. Unlike R_{-1}, There is no hope to decrease the gap between R_{-2} and R_+ by increasing the number of feedback bits as discussed in Section IV.

In Fig. 2, we have considered the setting where the main channel is a Rayleigh fading, whereas the eavesdropper's channel is a Rician fading described by [31]:

$$g = \sqrt{\frac{K}{K+1}}\,\bar{g} + \sqrt{\frac{1}{K+1}}\,g_w, \tag{37}$$

where K is the Rician factor, $\bar{g}$ is the mean component of g and g_w is the scattered (varying) component that follows

[3]The proof of this result is provided in Appendix J.

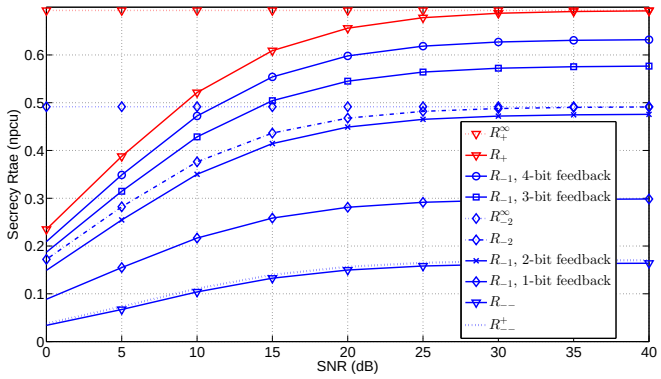

Figure 1. Achievable rates and the upper bound under STPC, for Rayleigh fading channels, with various q-bit feedback, $q = 1, 2, 3, 4$.

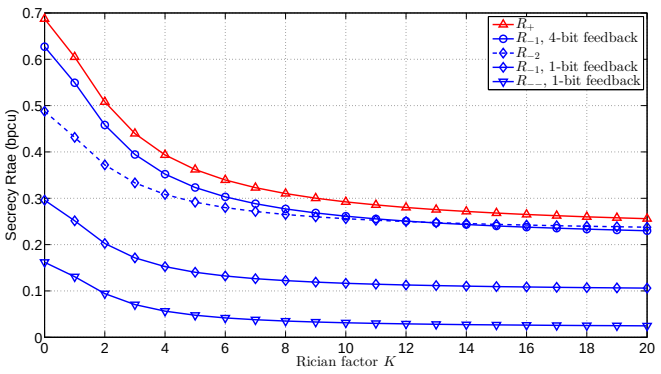

Figure 2. Achievable rates and the upper bound under STPC versus the Rician factor K, for q-bit feedback, $q = 1, 4$. The main channel is a normalized Rayleigh fading channel, whereas the eavesdropper's channel is a normalized Rician fading with factor K. The transmit power is equal to $P_{max} = 30$ dBs

a $CN(0, 1)$. This model captures the case where the eavesdropper in non-fading by setting $K \to \infty$, and captures the Rayleigh case by setting $K = 0$. For the model (37), the pdf

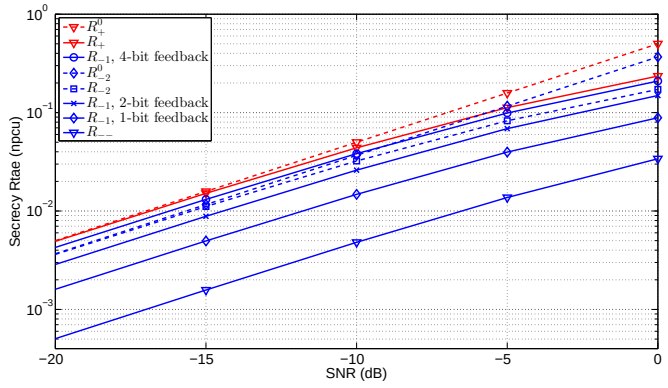

Figure 3. Achievable rates and the upper bound under STPC, for Rayleigh fading channels, with various q-bit feedback, $q = 1, 2, 4$, at low-SNR.

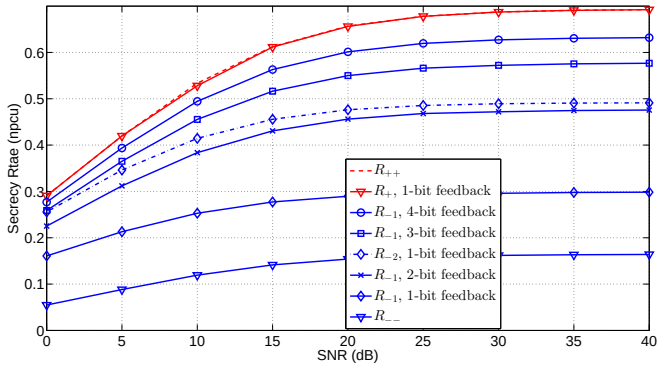

Figure 4. Achievable rates and the upper bound under LTPC, for Rayleigh fading channels, with various q-bit feedback, $q = 1, 2, 3, 4$.

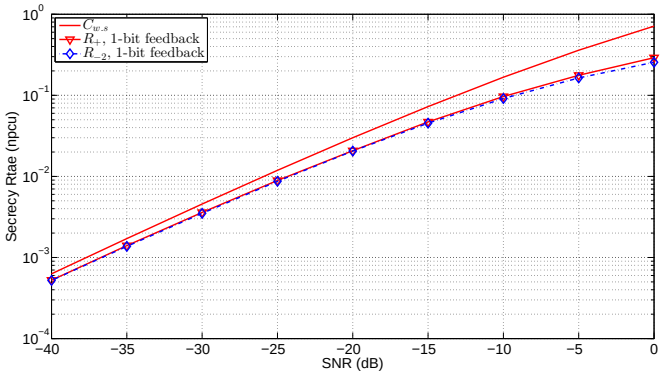

Figure 5. Achievable rate R_{-2} and the upper bound under LTPC, for Rayleigh fading channels, with 1-bit feedback, at low-SNR.

of γ_g is given by:

$$f_{\gamma_g}(\gamma_g) = (1+K)\ \exp\Big(-\big(1+K\big)*\gamma_g + K\Big)\Big)\ I_0\Big(\sqrt{4K\,(1+K)\,\gamma_g}\Big), \tag{38}$$

where $I_0(\cdot)$ is the Modified Bessel Function of the First Kind. The transmit power has been set to $P_{max} = 30$ dB to depict the high-SNR insight and we have considered a STPC. We have evaluated the rates R_{-1} and R_{-2} in Theorem 1 along with the upper bound R_+ for the proactive feedback (at the beginning of the blocks) scenario, versus the Rician factor K. In addition, we have evaluated the rate R_{--} presented in Theorem 3 for the ARQ feedback versus the Rician factor K. As shown in Fig.2, a positive secrecy rate is achievable even when the eavesdropper's channel is non fading ($K >> 1$) and even with an ARQ feedback. All rates decrease with K confirming that fading helps in providing secrecy. Note that with 4-bit feedback, the gap to the upper bound is roughly the same irrespective of the Rician factor K. However, differently from the Rayleigh case presented in the paper, R_{-2} with just 1-bit feedback outperforms R_{-1} with 4-bit feedback, for $K \geq 13$. This suggests that if the eavesdropper's channel is non-fading, then using a constant rate wiretap code is a better strategy than adapting the rate with the quantized main CSI feedback.

In Fig. 3, we have displayed performance of our lower

bounds and upper bound at low-SNR under STPC. The characterization in Corollary 3 is also reported in Fig. 3 where for convenience, we designate by R^0_{-2} and R^0_+ the RHS of (26) and (27), respectively. Although not fully characterized for a finite number of feedback bits, the secrecy capacity seems to scale linearly at low-SNR as shown in Fig. 3.

For LTPC, as shown in Fig. 4, performance has expectedly improved at finite SNR since one can exploit power adaptation in a more efficient way. However, at high-SNR, performances under STPC and LTPC are equal. We note that here again, with 4-bit feedback, more than 90% of the available capacity may be achieved by R_{-1}.

For the low-SNR regime, we have plotted in Fig. 5, the ergodic capacity of the main channel, the curves corresponding to the achievable rate R_{-2} and the upper bound R_+. We note first that the curves corresponding to R_{-2} and R_+ are undistinguishable for all SNR values below -10 dB, thus fully characterizing the capacity in this case. Furthermore, the three curves get closer as P_{max} tends toward zero, to actually coincide completely at $P_{max} \leq -70$ dBs (although not shown in Fig. 5), in full agreement with Theorem 4. While such low SNR values have a little practical meaning, the insight gained from our low SNR analysis seems to be very appealing.

VII. Conclusion

The secret message capacity of an ergodic block fading wiretap channels with limited-rate feedback has been addressed. Lower bounds and an upper bound have been derived when an arbitrary number of feedback bits at the beginning of each coherence block are provided to the sender by the legitimate receiver, through an error free public channel with limited capacity. We have also shown that a positive secrecy rate is achievable when only 1-bit ARQ feedback is given to the sender at the end of each coherence block. When the number of feedback bits is large enough, one of our lower bounds and the upper bounds coincide, thus fully characterizing the capacity in this case. While the capacity at high-SNR is bounded, it has been found that at asymptotically low-SNR regime, the lower and the upper bounds scale linearly with SNR under STPC whereas under LTPC and for a class of fading channels, 1-bit feedback is enough to achieve a secrecy rate equal to the ergodic capacity of the main channel as if there is no secrecy constraint. Our framework highlights the role of feedback in providing secure communication and emphasizes on the efficiency of secure communication at low-SNR regime as secrecy may be obtained with a marginal penalty.

We note that in this work, we have focused on a single layer coding approach; an interesting study would be to generalize our framework to a multi-layer coding or the so-called broadcast approach. Without secrecy constraint, It has been shown by previous studies that the broadcast approach outperforms single layer coding in terms of the average achievable rate, especially when the number of feedback bits is small (1 or 2). Whether this behavior holds or not under secrecy constraint is worth investigating.

References

[1] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[2] M. Bloch, J. Barros, M. Rodrigues, and S. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

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

[4] M. Bloch and J. Laneman, "Information-spectrum methods for information-theoretic security," in *Proc. Information Theory and Applications Workshop (ITA'2009)*, San Diego, CA, USA, Feb. 2009, pp. 23–28.

[5] Z. Rezki, A. Khisti, and M.-S. Alouini, "On the ergodic secret message capacity of the wiretap channel with finite-rate feedback," in *Proc. 2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Cambridge, MA, USA, july 2012, pp. 239 –243.

[6] M. Bloch and J. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE J. Select. Areas Commun.*, vol. 31, no. 9, pp. 1840–1849, Sep. 2013.

[7] Z. Rezki, A. Khisti, and M.-S. Alouini, "On the ergodic secrecy capacity of the wiretap channel under imperfect main channel estimation," in *Proc. 2011 45th Asilomar Conference on Signals, Systems and Computers (Asilomar'2011)*, Pacific Grove, CA, USA, Nov 2011, pp. 952–957.

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

[9] X. Zhou and M. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[10] A. Mukherjee and A. Swindlehurst, "Ensuring secrecy in MIMO wiretap channels with imperfect CSIT: A beamforming approach," in *2010 IEEE International Conference on Communications (ICC'2010)*, Cape Town, South Africa, May 2010, pp. 1–5.

[11] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, Jul. 2012.

[12] A. Mukherjee and A. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Processing*, vol. 59, no. 1, pp. 351–361, Jan 2011.

[13] W. Shi and J. Ritcey, "Robust beamforming for MISO wiretap channel by optimizing the worst-case secrecy capacity," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR'2010)*, Monterey, CA, USA, 7-10 Nov. 2010, pp. 300–304.

[14] J. Taylor, M. Hempel, H. Sharif, S. Ma, and Y. Yang, "Impact of channel estimation errors on effectiveness of eigenvector-based jamming for physical layer security in wireless networks," in *the 16th IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD'2011)*, Kyoto, Japan, Jun. 2011, pp. 122–126.

[15] S.-C. Lin, T.-H. Chang, Y.-L. Liang, Y. Hong, and C.-Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy with artificial noise: The noise leakage problem," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 901–915, March 2011.

[16] S. Gerbracht, C. Scheunert, and E. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inform. Forensics and Security*, vol. 7, no. 2, pp. 704–716, 2012.

[17] J. Li and A. Petropulu, "Explicit solution of worst-case secrecy rate for MISO wiretap channels with spherical uncertainty," *IEEE Trans. Signal Processing*, vol. 60, no. 7, pp. 3892–3895, 2012.

[18] S. Ma, M. Hong, E. Song, X. Wang, and D. Sun, "Outage constrained robust secure transmission for MISO wiretap channels," *submitted for publication, available at http://arxiv.org/pdf/1310.7158.pdf*, 2013.

[19] C.-W. Huang, T.-H. Chang, X. Zhou, and Y.-W. Hong, "Two-way training for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2724–2738, May 2013.

[20] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[21] T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, Apr. 2007.

[22] D. Love, R. W. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems,"

IEEE J. Select. Areas Commun., vol. 26, no. 8, pp. 1341–1365, october 2008.
[23] V. Hassel, D. Gesbert, M.-S. Alouini, and G. Oien, "A threshold-based channel state feedback algorithm for modern cellular systems," *IEEE Trans. Wireless Commun.,*, vol. 6, no. 7, pp. 2422–2426, july 2007.
[24] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
[25] A. Khisti and G. Wornell, "Secure transmission with multiple antennas. Part I: The MISOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.
[26] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," in *Proceedings of the IEEE International Symposium on Information Theory, Massachusetts, USA*, Aug. 1998, p. 42.
[27] Y. Abdallah, M. Latif, M. Youssef, A. Sultan, and H. El Gamal, "Keys through ARQ: Theory and practice," *IEEE Trans. Inform. Forensics and Security*, vol. 6, no. 3, pp. 737–751, Sep. 2011.
[28] I. Safaka, C. Fragouli, K. Argyraki, and S. Diggavi, "Creating shared secrets out of thin air," in *Proc. 11th ACM Workshop on Hot Topics in Networks, HotNets-XI*, Redmond, Seattle, WA, USA, Oct. 2012, pp. 73–78.
[29] R. Davidson and J.-Y. Duclos, "Testing for restricted stochastic dominance," *Econometric Reviews*, vol. 32, no. 1, pp. 84–125, 2013. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/07474938.2012.690332
[30] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Review*, vol. 44, no. 4, pp. 525–597, 2002.
[31] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, C. U. Press, Ed. Cambridge University Press, 2003.
[32] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, ser. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, July 2006.
[33] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.
[34] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
[35] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1986 –1992, Nov. 1997.

Appendix A
Proof of Theorem 1

A. Proof of the Lower Bound R_{-1}

Let $\{P_k\}_{k=1}^N$ be an arbitrary power policy in $\Pi^{(N)}$, and let $\{\tau_k\}_{k=1}^N$ be a family of reconstruction points in $\Theta^{(N)}$. We assume that the choice of rates $\{0 \le R_1 \le R_2, \ldots \le R_N < R_{N+1} = \infty\}$, where $R_p = \log\left(1 + P_p\, \tau_p\right)$, is selected in advance. Let $\Delta_p = \Pr(\tau_p \le \gamma_h < \tau_{p+1})$ for $p = 1, \ldots, N$. We establish that the rate $R_{-1} = \sum_{p=1}^N \Delta_p\, \mathbf{E}\left[R_p - \log(1+\gamma_g P_p)\right]^+ + \epsilon$ is achievable. We also let $R = \sum_{p=1}^N \Delta_p R_p - 2\epsilon$. We uniformly partition the set of all 2^{nR} sequences of length nR into $2^{nR_{-1}}$ bins so that there are $2^{n(R-R_{-1})}$ sequences per bin. Each message $W \in [1, 2^{nR_{-1}}]$ corresponds to one bin-index. To transmit a message W the transmitter selects the corresponding bin index and then select a binary sequence $\boldsymbol{v}$ uniformly at random from all of the sequences in that bin. Since all messages are equally likely, we induce a uniform distribution across all of 2^{nR} sequences. In each length m coherence block, we transmit the next $m \cdot R_p$ information bits using a Gaussian codebook. For convenience, we let the transmit codeword in coherence block i be $\boldsymbol{X}^m(i) = (X(i,1), \ldots, X(i,m))$ and the received sequences at the legitimate receiver and eavesdropper by $\boldsymbol{Y}^m(i) = (Y(i,1), \ldots, Y(i,m))$ and $\boldsymbol{Z}^m(i) = (Z(i,1), \ldots, Z(i,m))$, respectively. By weak law of large numbers, when L ($L \gg 1$) coherence periods are used for transmission, the entire binary sequence $\boldsymbol{v}$ is transmitted with high probability. Since in each block $R_p \le \log(1+\gamma_h P_p)$ holds, the receiver can decode the sequence $\boldsymbol{v}$ with high probability. For the secrecy analysis, we observe that from our construction the codeword sequence $\boldsymbol{X}^m(1), \boldsymbol{X}^m(2), \ldots \boldsymbol{X}^m(L)$ is independent and hence

$$H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L\right) = \sum_{i=1}^L H\left(\boldsymbol{X}^m(i) \mid \boldsymbol{Z}^m(i), h(i), g(i)\right) \quad (39)$$

Furthermore from the analysis of a Gaussian wiretap code we have that

$$H\left(\boldsymbol{X}^m(i) \mid \boldsymbol{Z}^m(i), h(i), g(i)\right) \ge m\, \mathbf{E}\left[R^{(i)} - \log\left(1+\gamma_g P^{(i)}\right) - \epsilon\right]^+ \quad (40)$$

where $R^{(i)} \in \{0, R_1, \ldots, R_N\}$ and $P^{(i)} \in \{0, P_1, \ldots, P_N\}$ are the rate and the power selected in block i. Thus, we have:

$$\begin{aligned} nR_e &= H\left(W \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{G}^L\right) && (41)\\ &\ge I\left(W; \boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L\right) && (42)\\ &= H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L\right) - H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, W\right) && (43)\\ &\ge \sum_{i=1}^L m\, \mathbf{E}\left[R^{(i)} - \log\left(1+\gamma_g P^{(i)}\right)\right]^+ - H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, W\right) && (44)\end{aligned}$$

Using weak-law of large numbers it can be seen that

$$\frac{1}{L}\sum_{i=1}^L [R^{(i)} - \log\left(1+\gamma_g P^{(i)}\right)]^+ \overset{L\to\infty}{\longrightarrow} \sum_{p=0}^{N-1} \Delta_p \mathbf{E}\left[\left[R_p - \log\left(1+\gamma_g P_p\right)\right]^+\right].$$

Thus it only remains to show that the second term satisfies $H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, W\right) \le n\epsilon$. We argue that given the message W the eavesdropper can uniquely decode the sequence $\boldsymbol{v}$ and hence the codeword sequence $\boldsymbol{X}^n$. In coherence block i, the eavesdropper constructs a list $\mathcal{L}_i$ for all codeword sequences that are jointly typical with the received sequence $\boldsymbol{Z}^m(i)$. From standard typicality analysis there are a total of $2^{m\left[R^{(i)} - \log(1+\gamma_g(i)P^{(i)})\right]^+}$ such sequences. It then searches for a unique sequence in the set $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \cdot \times \cdot \mathcal{L}_L$ that belongs to the bin-index of W. The size of this set is: $|\mathcal{L}| = 2^{m\sum_{i=1}^L \left[R^{(i)} - \log\left(1+\gamma_g(i)P^{(i)}\right)\right]^+}$, which from the weak-law of large numbers, approaches $2^{n(R_{-1}-\epsilon)}$ as $L \to \infty$. Since each sequence in the set $\mathcal{L}$ belongs to the bin of W with probability $2^{-nR_{-1}}$ the overall error probability can be shown to vanish as $n \to \infty$. Indeed, by the asymptotic equipartition property (AEP) and the Packing Lemma [32, Chap. 3], it can be shown that the probability of error at Eve is upper-bounded as follows:

$$\begin{aligned} P_e^{(Eve)} &\le \epsilon_1 + 2^{-nR_{-1}} \cdot 2^{m\sum_{i=1}^L \left[R^{(i)} - \log\left(1+\gamma_g(i)P^{(i)}\right)\right]^+}\\ &= \epsilon_1 + 2^{-nR_{-1}} \cdot 2^{n(R_{-1}-\epsilon)} && (45)\\ &= \epsilon_1 + 2^{-n\epsilon} && (46)\end{aligned}$$

where $\epsilon_1 \to 0$ as $n \to \infty$ and where (45) holds as $L \to \infty$. The right hand side of (46) vanishes to zero as $n \to \infty$ and hence so does $H\left(\boldsymbol{X}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, W\right)$ due to Fano's inequality.

B. Proof of the Lower Bound R_{-2}

We can think of the feedback as a deterministic mapping, say $\kappa(\cdot)$, such that $\kappa(\gamma_h) = k$ if $\gamma_h \in [\tau_k, \tau_{k+1})$. Then, we construct a new channel where the main channel fading is

amplified by $\sqrt{P(\kappa(\gamma_h))}$, i.e., the output $Y(i,j)$ in (1) becomes $\tilde{Y}(i,j) = \tilde{h}(i)\ X(i,j) + U(i,j)$, where $\tilde{h}(i) = \sqrt{P(\kappa(\gamma_h(i)))}\, h(i)$. Clearly, this is a specific use of CSI-T and thus the secrecy capacity of the new channel is not higher than the original one. Moreover, the new channel has no CSI-T and perfect CSI-R at the legitimate receiver. The rate R_{-2} follows then from [24, Corollary 2], by taking $V = X$ such that $p(x) = \mathcal{CN}(0,1)$. With this choice, the rate $I\left(X;\tilde{Y},\tilde{h}\right) - I\left(X;Z,\tilde{h},g\right) = I\left(X;\tilde{Y} \mid \tilde{h}\right) - I\left(X;Z,g \mid \tilde{h}\right)$ is achievable. The first term can be evaluated as follows:

$$I\left(X;\tilde{Y} \mid \tilde{h}\right) = \mathop{\mathbf{E}}_{\gamma_h}[\log(1 + P(\kappa(\gamma_h))\ \gamma_h)] \quad (47)$$

$$= \sum_{k=1}^{N} \mathop{\mathbf{E}}_{\tau_k \le \gamma_h < \tau_{k+1}} [\log(1 + \gamma_h P_k)]. \quad (48)$$

The second term can be evaluated similarly so that the rate

$$\sum_{k=1}^{N} \mathop{\mathbf{E}}_{\tau_k \le \gamma_g < \tau_{k+1}} \left[\log\left(\frac{1+\gamma_h P_k}{1+\gamma_g P_k}\right)\right] \quad (49)$$

is achievable. Maximizing over all P_k's and τ_k's subject to the power constraint completes the proof.

Appendix B
Proof of the Upper Bound R_+

We assume that the transmitter has CSI $u_i = \kappa(h_i)$ at time instant i, whereas the legitimate receiver knows $\gamma_{h,i}$. We mainly follow the approach in [33] with proper adaptation to secrecy and upper bound the equivocation rate as follows.

$$n R_e = H\left(W \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L\right) \quad (50)$$

$$\le I\left(W; \boldsymbol{Y}^n \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{u}^L\right) + n\delta_n \quad (51)$$

$$= \sum_{i=1}^{n} I\left(W; Y_i \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{Y}^{i-1}, \boldsymbol{u}^L\right) + n\delta_n \quad (52)$$

$$= \sum_{i=1}^{n} \Big\{ h\left(Y_i \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{Y}^{i-1}, \boldsymbol{u}^L\right) - h\left(Y_i \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{Y}^{i-1}, \boldsymbol{u}^L, W\right) \Big\} + n\delta_n \quad (53)$$

$$\le \sum_{i=1}^{n} h\left(Y_i \mid Z_i, h_i, g_i, \boldsymbol{u}^L\right) - h\left(Y_i \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{Y}^{i-1}, \boldsymbol{u}^L, W, X_i\right) + n\delta_n \quad (54)$$

$$= \sum_{i=1}^{n} h\left(Y_i \mid Z_i, h_i, g_i, \boldsymbol{u}^L\right) - h\left(Y_i \mid Z_i, h_i, g_i, X_i, U^L\right) + n\delta_n \quad (55)$$

$$= \sum_{i=1}^{n} I\left(X_i; Y_i \mid Z_i, h_i, g_i, \boldsymbol{u}^L\right) + n\delta_n \quad (56)$$

$$\le \sum_{i=1}^{n} \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_{h,i} P_i\left(\boldsymbol{u}^i\right)}{1+\gamma_{g,i} P_i\left(\boldsymbol{u}^i\right)}\right)\right]^+\right] + n\delta_n, \quad (57)$$

where (51) follows from Fano's inequality and also because $\boldsymbol{u}^L$ is a deterministic function of $\boldsymbol{h}^L$ and where (57) follows because given h_i and g_i, the channel at hand is a wiretap channel with average transmit power $P_i\left(\boldsymbol{u}^i\right)$, where $P_i\left(\boldsymbol{u}^i\right) = \mathbf{E}\left[|X_i|^2 \,\middle|\, \boldsymbol{u}^i\right]$, since given $\boldsymbol{u}^L$, X_i is independent of h_i. The above upper bound is tight if $\boldsymbol{X}^n$ is a sequence with zero-mean Gaussian components X_i, statistically independent conditionally on $\boldsymbol{u}^L$. Let $X_i = \sqrt{P_i(\boldsymbol{u}^i)}\, T_i$, where T_i is i.i.d. $\mathcal{CN}(0,1)$. Then we need only to prove that the above upper bound is maximized by a power allocation $P_i\left(\boldsymbol{u}^i\right) = \lambda(u_i)$, a time-invariant function of u_i only. To do this, we have:

$$\mathbf{E}\left[\left[\log\left(\frac{1+\gamma_{h,i} P_i\left(\boldsymbol{u}^i\right)}{1+\gamma_{g,i} P_i\left(\boldsymbol{u}^i\right)}\right)\right]^+\right]$$

$$= \mathbf{E}\left[\mathbf{E}\left[\left[\log\left(\frac{1+\gamma_{h,i} P_i\left(\boldsymbol{u}^i\right)}{1+\gamma_{g,i} P_i\left(\boldsymbol{u}^i\right)}\right)\right]^+ \,\middle|\, \gamma_{h,i}, \gamma_{g,i}, u_i\right]\right] \quad (58)$$

$$\le \mathbf{E}\left[\left[\log\left(\frac{1+\mathbf{E}\left[\gamma_{h,i} P_i\left(\boldsymbol{u}^i\right) \,\middle|\, \gamma_{h,i}, \gamma_{g,i}, u_i\right]}{1+\mathbf{E}\left[\gamma_{g,i} P_i\left(\boldsymbol{u}^i\right) \,\middle|\, \gamma_{h,i}, \gamma_{g,i}, u_i\right]}\right)\right]^+\right] \quad (59)$$

$$= \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_{h,i} \mathbf{E}\left[P_i\left(\boldsymbol{u}^i\right) \,\middle|\, u_i\right]}{1+\gamma_{g,i} \mathbf{E}\left[P_i\left(\boldsymbol{u}^i\right) \,\middle|\, u_i\right]}\right)\right]^+\right] \quad (60)$$

$$= \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_{h,i} \lambda_i(u_i)}{1+\gamma_{g,i} \lambda_i(u_i)}\right)\right]^+\right] \quad (61)$$

$$= \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_h \lambda_i(u)}{1+\gamma_g \lambda_i(u)}\right)\right]^+\right] \quad (62)$$

where (59) follows from Jensen's inequality since the function $x \mapsto \left[\log\left(\frac{1+ax}{1+bx}\right)\right]^+$ is concave for any positive a and b; where (60) follows because conditioned on u_i, $\boldsymbol{u}^i$ is independent of $\gamma_{h,i}$ and $\gamma_{g,i}$ due to the fact that the fading process $\{\gamma_{h,i}\}$ is i.i.d.; where we have defined $\lambda_i(u_i)$ in (61) as $\lambda_i(u_i) = \mathbf{E}\left[P_i\left(\boldsymbol{u}^i\right) \,\middle|\, u_i\right]$. Since the fading processes $\{h_i\}$ and $\{g_i\}$ are ergodic and stationary, then they have a stationary first-order distribution and thus the expectation in (61) does not depend on their time index i, from which (62) follows. Combining (57) and (62), we obtain:

$$R_e \le \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_h \lambda_i(u)}{1+\gamma_g \lambda_i(u)}\right)\right]^+\right] + \delta_n \quad (63)$$

$$\le \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_h \frac{1}{n}\sum_{i=1}^{n}\lambda_i(u)}{1+\gamma_g \frac{1}{n}\sum_{i=1}^{n}\lambda_i(u)}\right)\right]^+\right] + \delta_n \quad (64)$$

$$= \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_h \lambda(u)}{1+\gamma_g \lambda(u)}\right)\right]^+\right] + \delta_n \quad (65)$$

where (64) follows again by Jensen's inequality and where $\lambda(u)$ in (65) is defined as $\lambda(u) = \frac{1}{n}\sum_{i=1}^{n}\lambda_i(u)$. The above upper bound is tight if $\lambda_i(u)$ is independent of i. Letting $n \to \infty$ and maximizing over all power policies $\{\lambda(u)\}$ that satisfy the STPC (resp. LTPC), we establish that

$$R_e \le \max_{\lambda(u)\ \text{s.t.}\ \text{STPC}} \mathbf{E}\left[\left[\log\left(\frac{1+\gamma_h \lambda(u)}{1+\gamma_g \lambda(u)}\right)\right]^+\right]. \quad (66)$$

Since $u = \kappa(h)$, where $\kappa(\cdot)$ is a deterministic mapping, the upper bound in Theorem 2 follows. It remains to show that the lower and the upper bounds coincide as $N \to \infty$. For this purpose, let us choose τ_k's such that $\left(F_{\gamma_h}(\tau_{k+1}) - F_{\gamma_h}(\tau_k)\right) = \frac{1}{N}$. Note that this is possible as long as Alice is aware of the

statistics of the main channel gain h which is the case. The results follows then as $N \to \infty$ due to the ergodicity.

Appendix C
Proof of Theorem 3

The achievability scheme is similar to that of Theorem 1, with the difference that because the sender keeps repeating the blocks that are NACKed until she receives an ACK, these repetitions leak additional information to the eavesdropper. Again the Random Coding Theorem ensures that there exists a Gaussian codebook of rate $R = \log(1+\tau P)$ such that the fraction of successfully decoded frames is given by:

$$\Pr(\text{success}) = 1 - \Pr\{R > \log(1+\gamma_h P)\}. \tag{67}$$

For the secrecy analysis, we first let L_i be the number of blocks that have been repeated i times, $i = 0, \ldots, \infty$. For instance, L_0 represents the number of blocks that have not been repeated, L_1 represents the number of blocks that have been repeated once and so on. Also, let s be a binary random variable that describes the ARQ feedback. That is, $s = 1$ if an ACK is received and $s = 0$ otherwise. One can upper bound the equivocation rate as follows:

$$\begin{aligned}
nR_e &= H\left(W \mid \boldsymbol{Z}^n, h^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (68)\\
&\geq I\left(W; \boldsymbol{X}^{mL_0} \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (69)\\
&= h\left(\boldsymbol{X}^{mL_0} \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) \\
&\quad - h\left(\boldsymbol{X}^{mL_0} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (70)\\
&= h\left(\boldsymbol{X}^{mL_0} \mid \boldsymbol{Z}^{mL_0}, \boldsymbol{h}^{L_0}, \boldsymbol{g}^{L_0}\right) \\
&\quad - h\left(\boldsymbol{X}^{mL_0} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (71)\\
&\geq \sum_{i=1}^{L_0} m\left\{\left[R - \epsilon - \log\left(1+\gamma_g(i)\,P\right)\right]^+\right\} \\
&\quad - h\left(\boldsymbol{X}^{mL_0} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right), && (72)
\end{aligned}$$

where (71) follows because the eavesdropper does not gain any information about X^{mL_0} by observing the remaining $L - L_0$ blocks, since the blocks are independents and the channel is memoryless. The second term on the RHS of (72) can be made arbitrary small using a list decoding argument similarly to Appendix A. Finally, when $L_0 \to \infty$, the ratio $\frac{L_0}{L}$ can be computed as follows:

$$\begin{aligned}
\lim_{L_0\to\infty} \frac{L_0}{L} &= \lim_{L_0\to\infty} \frac{1}{L} \sum_{i=1}^{L} \mathbb{1}_i && (73)\\
&= \Pr\{\text{no repetition}\} \\
&= \Pr\{\text{blocks } i \text{ and } (i-1) \text{ not repeated}\} && (74)\\
&= \Pr(\text{success})^2 && (75)\\
&= \theta^2, && (76)
\end{aligned}$$

where $\mathbb{1}_i$ is an indicator function that is equal to one if the block i is not repeated and is equal to zero otherwise. Using the ergodicity of the channel in (72) along with (76), (8) follows immediately.

Appendix D
Proof of Corollary 1

The existence of a codebook with arbitrary low error probability is justified similarly as in Appendix C. Here, we outline the secrecy analysis.

$$\begin{aligned}
nR_e &\geq h\left(\boldsymbol{X}^{mL_0}, \boldsymbol{X}^{mL_1} \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) \\
&\quad - h\left(\boldsymbol{X}^{mL_0}, \boldsymbol{X}^{mL_1} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (77)\\
&= h\left(\boldsymbol{X}^{mL_0} \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) \\
&\quad + h\left(\boldsymbol{X}^{mL_1} \mid \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L, \boldsymbol{X}^{mL_0}\right) \\
&\quad - h\left(\boldsymbol{X}^{mL_0}, \boldsymbol{X}^{mL_1} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (78)\\
&= h\left(\boldsymbol{X}^{mL_0} \mid \boldsymbol{Z}^{mL_0}, \boldsymbol{h}^{L_0}, \boldsymbol{g}^{L_0}\right) \\
&\quad + h\left(\boldsymbol{X}^{mL_1} \mid \boldsymbol{Z}^{2mL_1}, \boldsymbol{h}^{2L_1}, \boldsymbol{g}^{2L_1}\right) \\
&\quad - h\left(\boldsymbol{X}^{mL_0}, \boldsymbol{X}^{mL_1} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right) && (79)\\
&\geq \sum_{i=1}^{L_0} m\left\{\left[R - \epsilon - \log\left(1+\gamma_g(i)\,P\right)\right]^+\right\} \\
&\quad + \sum_{i=1}^{L_1} m\left\{\left[R - \epsilon - \log\left(1+\gamma_g^{(2)}(i)\,P\right)\right]^+\right\} \\
&\quad - h\left(\boldsymbol{X}^{mL_0}, \boldsymbol{X}^{mL_1} \mid W, \boldsymbol{Z}^n, \boldsymbol{h}^L, \boldsymbol{g}^L, \boldsymbol{s}^L\right), && (80)
\end{aligned}$$

where (79) follows because the eavesdropper does not gain any information about $\boldsymbol{X}^{mL_0}$ and $\boldsymbol{X}^{mL_1}$ by observing the remaining $(L - L_0)$ and $(L - L_1)$ blocks, respectively, since the blocks are independents and the channel is memoryless. To obtain (80), we expand the first term in (79) exactly as in the case of no repetition (please see Appendix C) whereas the second term in (79) can be expanded as follows:

$$\begin{aligned}
&h\left(\boldsymbol{X}^{mL_1} \mid \boldsymbol{Z}^{2mL_1}, \boldsymbol{h}^{2L_1}, \boldsymbol{g}^{2L_1}\right) \\
&= \sum_{\substack{\text{blocks } i \\ \text{repeated once}}} h\left(\boldsymbol{X}^m(i) \mid \boldsymbol{Z}^m(i), \boldsymbol{Z}^m(i+1), h(i), h(i+1)\right. \\
&\quad \left., g(i), g(i+1)\right) && (81)\\
&= \sum_{\substack{\text{blocks } i \\ \text{repeated once}}} \left[h\left(\boldsymbol{X}^m(i)\right)\right. \\
&\quad \left. - I\left(\boldsymbol{X}^m(i); \boldsymbol{Z}^m(i), \boldsymbol{Z}^m(i+1), h(i), h(i+1), g(i), g(i+1)\right)\right]^+ && (82)\\
&= \sum_{\substack{\text{blocks } i \\ \text{repeated once}}} \left[h\left(\boldsymbol{X}^m(i)\right)\right. \\
&\quad \left. - I\left(\boldsymbol{X}^m(i); \boldsymbol{Z}^m(i), \boldsymbol{Z}^m(i+1) \mid h(i), h(i+1), g(i), g(i+1)\right)\right]^+ && (83)\\
&\geq \sum_{i=1}^{L_1} \left\{m\left[R - \epsilon - \log\left(1+\gamma_g^{(2)} P\right)\right]^+\right\}, && (84)
\end{aligned}$$

where (81) follows again from the independence of the block pairs that have been repeated once and from the construction of the codeword sequence $X(1), \ldots, X(L)$, and where (83) follows because $X(i)$ and $(H(i), H(i+1), G(i), G(i+1))$ are independent and where (84) follows from the fact that Gaussian random variables are entropy maximizers. The third term on the RHS of (80) can be made arbitrary small using a list decoding argument similarly to Appendix C. Finally, when $L_0 \to \infty$ and $L_1 \to \infty$, the ratios $\frac{L_0}{L}$ converges to θ^2 due to

(76), whereas $\frac{L_1}{L}$ can be computed as follows:

$$\lim_{L_1\to\infty}\frac{L_1}{L} = \Pr\{\text{blocks } i \text{ and } (i-2) \text{ are not repeated and } (i-1) \text{ repeated}\} \quad (85)$$

$$= \theta^2\ (1-\theta). \quad (86)$$

Using the ergodicity of the channels in (80) along with (76) and (86), the equivocation rate can be upper-bounded by:

$$R_e \geq \theta^2 \underset{\gamma_g}{\mathbf{E}}\left[\left[R-\log\left(1+\gamma_g P\right)\right]^+\right] + \theta^2\ (1-\theta) \underset{\gamma_g^{(2)}}{\mathbf{E}}\left[\left[R-\log\left(1+\gamma_g^{(2)} P\right)\right]^+\right],$$

from which the result in Corollary 1 follows immediately.

Appendix E
Proof of Lemma 1

For brevity, we prove Lemma 1 for $N=1$ and $N=2$, after that, it becomes clear that generalization of the proof to arbitrary N follows immediately.

1) $N=1$

In this case. the rate R_{-2} may be written as $R_{-2} = \max_{\substack{0\leq P\leq P_{max}\\ 0\leq\tau}} \underset{\substack{\tau\leq\gamma_h\\ \gamma_g}}{\mathbf{E}}\left[\log\left(\frac{1+\gamma_h P}{1+\gamma_g P}\right)\right]$. Writing the KKT condition with respect to τ, the optimal values P^* and τ^* must satisfy:

$$f_{\gamma_h}(\tau^*)\left(\log(1+\tau^* P^*) - \mathbf{E}\left[\log\left(1+\gamma_g P\right)\right]\right) = 0. \quad (87)$$

Since we focus on positive PDF's that can be null only at $x=0$, we conclude that $\tau^* = \frac{e^{C(P^*)}-1}{P^*}$, where $C(P) = \mathbf{E}\left[\log\left(1+\gamma_g P\right)\right]$, is the maximizer of $K(\tau, P^*)$. Note that $\tau^* > 0$, for any $P^* > 0$.

To show that $P^* = P_{max}$, let us define the function $K(\tau, P)$ over $[0,\infty)\times[0, P_{max}]$ by: $K(\tau,P) = \underset{\substack{\tau\leq\gamma_h\\ \gamma_g}}{\mathbf{E}}\left[\log\left(\frac{1+\gamma_h P}{1+\gamma_g P}\right)\right]$. Our objective is to show that $K(\tau,P)$ is increasing in P and hence setting $P^* = P_{max}$ can only increase the objective function. For this purpose, we have:

$$K(\tau,P) = \int_\tau^\infty \log(1+xP)\ f_{\gamma_h}(x)\ \mathrm{d}x - \left(1-F_{\gamma_h}(\tau)\right)\int_0^\infty \log(1+xP)\ f_{\gamma_g}(x)\ \mathrm{d}x \quad (88)$$

$$= \left(1-F_{\gamma_h}(\tau)\right)\int_0^\infty \log(1+xP)\left[\frac{f_{\gamma_h}(x)\ \mathbb{1}_{[\tau,\infty)}(x)}{1-F_{\gamma_h}(\tau)} - f_{\gamma_g}(x)\right]\mathrm{d}x \quad (89)$$

$$= \left(1-F_{\gamma_h}(\tau)\right)\left[\mathbf{E}\left[\log\left(1+\gamma_{h,[\tau,\infty)}P\right)\right] - \mathbf{E}\left[\log\left(1+\gamma_g P\right)\right]\right] \quad (90)$$

where (90) follows because the function $x\mapsto\frac{f_{\gamma_h}(x)\mathbb{1}_{[\tau,\infty)}(x)}{1-F_{\gamma_h}(\tau)}$ is the PDF of the r.v. $\gamma_{h,[\tau,\infty)} = \gamma_h \mid \gamma_h \geq \tau$. Using (90), the derivative of $K(\tau,P)$ can written as

$$\frac{\partial K(\tau,P)}{\partial P} = \left(1-F_{\gamma_h}(\tau)\right)\left[\mathbf{E}\left[\frac{\gamma_{h,[\tau,\infty)}}{1+\gamma_{h,[\tau,\infty)}P}\right] - \mathbf{E}\left[\frac{\gamma_g}{1+\gamma_g P}\right]\right]. \quad (91)$$

Now, we need the following facts which are know results in stochastic dominance theory.

Fact 1: If $X \geq Y$, then $\mu(X) \geq \mu(Y)$, for any increasing mapping $\mu(\cdot)$.

as a corollary of Fact 1, we also have:

Fact 2: If $X \geq Y$, then $\mathbf{E}[X] \geq \mathbf{E}[Y]$.

By assumption of Lemma 1, we know that $\gamma_{h,[\tau,\infty)} \geq \gamma_g$ and since $x\mapsto\frac{x}{1+xP}$ is an increasing mapping, then by Fact 1 and Fact 2, we have $\frac{\partial K(\tau,P)}{\partial P} \geq 0$ and hence $K(\tau,P)$ is increasing in P. Therefore $P^* = P_{max}$ is optimal. In addition, since $K(\tau,0)=0$, then R_{-2} is necessarily non-negative. Lemma 1 is thus proved for $N=1$.

1) $N=2$

In this case. the rate R_{-2} may be written as

$$R_{-2} = \max_{\substack{0\leq P_1\leq P_{max}\\ 0\leq P_2\leq P_{max}\\ 0\leq\tau_1\leq\tau_2}} \left\{ \underset{\substack{\tau_1\leq\gamma_h\leq\tau_2\\ \gamma_g}}{\mathbf{E}}\left[\log\left(\frac{1+\gamma_h P_1}{1+\gamma_g P_1}\right)\right] + \underset{\substack{\tau_2\leq\gamma_h\\ \gamma_g}}{\mathbf{E}}\left[\log\left(\frac{1+\gamma_h P_2}{1+\gamma_g P_2}\right)\right]\right\}. \quad (92)$$

Again, the KKT condition with respect to τ_1 implies that

$$\tau_1^* = \frac{e^{C(P_1^*)}-1}{P_1^*}, \quad (93)$$

whereas with respect to τ_2 gives the necessary condition

$$\frac{1+\tau_2^* P_1^*}{1+\tau_2^* P_2^*} = \frac{e^{C(P_1^*)}}{e^{C(P_2^*)}}. \quad (94)$$

Now, assume that $P_1^* < P_2^*$. Let us define the function $F(\cdot)$ on $[0,\infty)$ by: $F(\tau) = \frac{1+\tau P_1^*}{1+\tau P_2^*} - \frac{e^{C(P_1^*)}}{e^{C(P_2^*)}}$. One can easily verify that $F(\cdot)$ is strictly monotonically decreasing, that $F(0) = 1-\frac{e^{C(P_1^*)}}{e^{C(P_2^*)}} > 0$ since $P_1^* < P_2^*$ and that $\lim_{\tau\to\infty} F(\tau) = \frac{P_1^*}{P_2^*} - \frac{e^{C(P_1^*)}}{e^{C(P_2^*)}}$. Note that

$$\frac{e^{C(P_1^*)}}{e^{C(P_2^*)}} = \exp\left(\mathbf{E}\left[\log\left(\frac{1+g P_1^*}{1+g P_2^*}\right)\right]\right) \quad (95)$$

$$> \exp\left(\mathbf{E}\left[\log\left(\frac{P_1^*}{P_2^*}\right)\right]\right) \quad (96)$$

$$= \frac{P_1^*}{P_2^*}. \quad (97)$$

Thus, $\lim_{\tau\to\infty} F(\tau) < 0$. Since $F(\cdot)$ is strictly monotonically decreasing, then there should exist a unique $\tau_2^* > 0$, such that $F\left(\tau_2^*\right) = 0$, and since $\tau_1^* \leq \tau_2^*$, then either $\tau_1^* = \tau_2^*$ or $F\left(\tau_1^*\right) > 0$. We rule out the last condition so that we are left with the necessary condition $\tau_1^* = \tau_2^*$. Taking into account (93), the condition $F\left(\tau_1^*\right) > 0$ can be equivalently written as

$$\frac{e^{C(P_1^*)}-1}{P_1^*} < \frac{e^{C(P_2^*)}-1}{P_2^*}. \quad (98)$$

Let $G(\cdot)$ the function defined on $(0,\infty)$ by $G(P) = \frac{e^{C(P)}-1}{P}$. Its derivative, denoted $G'(\cdot)$ can be computed as $G'(P) = -\frac{e^{C(P)}}{P^2}\left(1-PC'(P)-e^{-C(P)}\right)$. We show that $G'(P) \leq 0$ for all $P\in(0,\infty)$ as follows:

$$1-PC'(P) = \mathbf{E}\left[\frac{1}{1+gP}\right] \quad (99)$$

$$= e^{\log\left(\mathbf{E}\left[\frac{1}{1+gP}\right]\right)} \quad (100)$$

$$> e^{\mathbf{E}\left[\log\left(\frac{1}{1+gP}\right)\right]} \quad (101)$$

$$= e^{-C(P)}, \quad (102)$$

where (101) follows by the Jensen's inequality since $\log(\cdot)$ is strictly concave. Hence $G(\cdot)$ is strictly monotonically decreasing on $(0,\infty)$ and since $P_1^* < P_2^*$, then we have $G\left(P_1^*\right) > G\left(P_2^*\right)$. Thus condition (98) cannot hold and neither can $F\left(\tau_1^*\right) > 0$. Therefore, we must have $\tau_1^* = \tau_2^*$ if $P_1^* < P_2^*$. But, if this is the case, the first part of the rate R_{-2} contributes nothing to the objective function. Also, if $P_1^* = P_2^*$, the first part can be merged with the second part so that R_{-2} can be written simply as:

$$R_{-2} = \max_{\substack{0\le P\le P_{max}\\ 0\le \tau}} \mathop{\mathbf{E}}_{\tau\le\frac{\gamma_h}{\gamma_g}} \left[\log\left(\frac{1+\gamma_h P}{1+\gamma_g P}\right)\right], \quad (103)$$

which corresponds exactly to the case $N = 1$. This completes the proof of Lemma 1.

Appendix F
On the positiveness of μ for $\mathcal{P}_2$

We assume here that the main and the eavesdropper channels are identically distributed so that $f_{\gamma_h}(x) = f_{\gamma_g}(x)$ for all x in the support of γ_h and γ_g. Let us write the $\frac{\partial}{\partial P_k}$ condition in Table I for $k = N$ as:

$$\frac{1}{1-F_{\gamma_h}(\tau_N)} \int_{\tau_N}^{\infty} \frac{\gamma_h}{1+\gamma_h P_N} f_{\gamma_h(\gamma_h)}\, \mathrm{d}\gamma_h - \int_0^{\infty} \frac{\gamma_g}{1+\gamma_g P_N} f_{\gamma_h(\gamma_g)}\, \mathrm{d}\gamma_g = \mu. \quad (104)$$

Let us consider a function $K(\cdot)$ defined on $[0,\infty)$ by: $K(x) = \frac{1}{1-F_{\gamma_h}(x)} \int_x^{\infty} \frac{\gamma_h}{1+\gamma_h P_N} f_{\gamma_h(\gamma_h)}\, \mathrm{d}\gamma_h$. The function $K(\cdot)$ is monotonically increasing since its derivative $K'(\cdot)$ given by

$$K'(x) = \frac{f_{\gamma_h}(x)}{\left(1-F_{\gamma_h}(x)\right)^2} \int_x^{\infty} \left(\frac{\gamma_h}{1+\gamma_h P_N} - \frac{x}{1+x P_N}\right) f_{\gamma_h(\gamma_h)}\, \mathrm{d}\gamma_h \quad (105)$$

is non-negative. If in addition $f_{\gamma_h}(x) > 0$ for all $x > 0$, then since the integrand in (105) is strictly positive for for all $\gamma_h \in (x,\infty)$, then $K'(x) > 0$ on $(0,\infty)$ and thus $K(\cdot)$ is strictly increasing on $(0,\infty)$. Our claim is that $\forall x > 0$, $K(x) > K(0)$. Assume that $\exists\, x_0 > 0$ such that $K(x_0) = K(0)$. Take any $\epsilon > 0$ such that $x_0 - \epsilon > 0$, we have that $K(x_0-\epsilon) < K(x_0) = K(0)$ due the monotonicity of $K(\cdot)$. But since $K(\cdot)$ is a continuous function, the later statement implies that $\exists\, \eta \in (0, x_0-\epsilon)$ such that $K(x_0-\epsilon) < K(\eta) < K(0)$. Note that the most left side of the later inequality contradicts the fact that $K(\cdot)$ is strictly increasing on $(0,\infty)$. Therefore, our claim holds true and since τ_N is positive, we have $K(\tau_N) > K(0)$ so that μ in (104) is positive. In summary, to prove that μ is positive, we needed to assume that γ_h and γ_g are identically distributed and that $f_{\gamma_h}(x) > 0$ for all $x > 0$.

Appendix G
Proof of Corollary 2

We prove (20) by showing that $\lim_{P_{max}\to\infty} R_{-1} = R_{-1}^{\infty}$. Let $\tilde{R}_{-1}$ be defined by:

$$\tilde{R}_{-1} = \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g} \left[\left[\log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}}\right)\right]^+\right],$$

for some τ_k's, $\tau_1 \le \tau_2 \le \ldots \le \tau_N$. Since for any P_{max} value,

$$\left[\log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}}\right)\right]^+ \le \left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+,$$

for all τ_k and all γ_g; and since

$$\mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right] < \infty;$$

due to the fact that $f_{\gamma_g}(\cdot)$ is continuous and bounded and $\left|\int_0^1 \log(x)\, dx\right| = 1$; then by the Dominated Convergence Theorem, we have:

$$\lim_{P_{max}\to\infty} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}}\right)\right]^+\right] = \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\lim_{P_{max}\to\infty} \log\left(\frac{1+\tau_k P_{max}}{1+\gamma_g P_{max}}\right)\right]^+\right] = \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right], \quad (106)$$

which implies that $\lim_{P_{max}\to\infty} \tilde{R}_{-1} = \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right]$. Hence, $\forall \epsilon > 0$, there exists a certain P_0 such that $\forall P_{max} > P_0$, we have:

$$\left|\tilde{R}_{-1} - \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\log\left(\frac{\tau_k}{\gamma_g}\right)\right]^+\right]\right| \le \epsilon. \quad (107)$$

Taking the maximum over all τ_k's such that $\tau_1 \le \tau_2 \le \ldots \le \tau_N$ on both sides of (107) establishes the desired result. The proof of (21) follows along similar lines. To prove (22), we simply recall that under STPC, $R_+ = R_{++}$ and that $\lim_{P_{max}\to\infty} R_{++} = \mathop{\mathbf{E}}_{\gamma_h,\gamma_g}\left[\left[\log\left(\frac{\gamma_h}{\gamma_g}\right)\right]^+\right]$ as it has been proved in [3], thus $\lim_{P_{max}\to\infty} R_+ = R_{++}^{\infty}$.

Appendix H
Proof of Corollary 3

In order to prove (25), we proceed similarly as in the proof of Corollary 2 and verify that

$$\lim_{P_{max}\to 0} \frac{\tilde{R}_{-1}}{P_{max} \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k - \gamma_g\right]^+\right]} = 1.$$

This implies that $\forall\, \epsilon > 0$, $\exists\, \eta > 0$ such that if $P_{max} \le \eta$, then $\left|\frac{\tilde{R}_{-1}}{P_{max}\sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathbf{E}_{\gamma_g}\left[\left[\tau_k-\gamma_g\right]^+\right]} - 1\right| \le \epsilon$. This, in turn, implies that:

$$(1-\epsilon) P_{max} \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k - \gamma_g\right]^+\right] < \tilde{R}_{-1} < (1+\epsilon) P_{max} \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k - \gamma_g\right]^+\right] \quad (108)$$

Taking the maximum over all τ_k's, $\tau_1 \le \ldots \le \tau_N$, on both sides of the last two inequalities, we obtain:

$$\left|\frac{R_{-1}}{P_{max} \cdot \max_{0\le\tau_1\le\ldots\le\tau_N} \sum_{k=1}^{N} \Pr\{\tau_k \le \gamma_h < \tau_{k+1}\} \cdot \mathop{\mathbf{E}}_{\gamma_g}\left[\left[\tau_k - \gamma_g\right]^+\right]} - 1\right| < \epsilon, \quad (109)$$

which proves (25). The proof of (26) follows along similar lines, and thus one can prove that:

$$\left| \frac{R_{-2}}{P_{max} \cdot \max\limits_{\tau \geq 0} \mathop{\mathbf{E}}\limits_{\gamma_h \geq \tau, \gamma_g} \left[\gamma_h - \gamma_g\right]} - 1 \right| < \epsilon. \tag{110}$$

In addition, $\tau^* = \mathop{\mathbf{E}}\limits_{\gamma_g}\left[\gamma_g\right]$ is the maximizer of the denominator in (110) and hence the proof of (26) is completed. Finally, the proof of (27) follows from a series expansion of $\log\left[\frac{1+P_{max}\,\gamma_h}{1+P_{max}\,\gamma_g}\right]$ around $P_{max} = 0$ to the second order and by averaging the obtained expression.

Appendix I
Proof of Theorem 4

Since the capacity without secrecy constraint cannot be smaller than the one under secrecy constraint, the converse part of Theorem 4 is immediate. To prove the achievability part, let us first define the maximum channel gain G by [34]: $G = \sup\limits_{p(x)} \frac{\mathbf{E}\left[\gamma_h |x|^2\right]}{\mathbf{E}\left[|x|^2\right]}$. Let us consider the conditional input distribution defined by

$$f_{x|\gamma_h}(x \mid \gamma_h) = \begin{cases} \delta\left(x - \sqrt{P_0}\right) & \text{if} \quad \gamma_h \geq \nu, \\ \delta(x) & \text{otherwise,} \end{cases} \tag{111}$$

where $\delta(\cdot)$ is the Dirac delta function, where $P_0 = \frac{P_{max}}{1-F_{\gamma_h}(\nu)}$ and where ν is a threshold that needs to be determined. Clearly, the input distribution (111) satisfies the LTPC since:

$$\begin{aligned}
\mathbf{E}\left[|x|^2\right] &= \int_{-\infty}^{+\infty} |x|^2 f_x(x)\, \mathrm{d}x && (112)\\
&= \int_{-\infty}^{+\infty} |x|^2 \left(F_{\gamma_h}(\nu)\,\delta(x) + \left(1 - F_{\gamma_h}(\nu)\right)\delta\left(x - \sqrt{P_0}\right)\right) \mathrm{d}x && (113)\\
&= \left(1 - F_{\gamma_h}(\nu)\right) P_0 && (114)\\
&= P_{max}. && (115)
\end{aligned}$$

Furthermore, we verify that:

$$\begin{aligned}
\lim_{P_{max}\to 0} \frac{|\mathbf{E}[x]|^2}{\mathbf{E}\left[|x|^2\right]} &= \lim_{P_{max}\to 0} \left(1 - F_{\gamma_h}(\nu)\right) && (116)\\
\lim_{P_{avg}\to 0} \frac{\mathbf{E}\left[\gamma_h |x|^2\right]}{\mathbf{E}\left[|x|^2\right]} &= \lim_{P_{max}\to 0} \frac{\int_{\nu}^{\infty} \gamma_h f_{\gamma_h}(\gamma_h)}{1 - F_{\gamma_h}(\nu)}. && (117)
\end{aligned}$$

Now, choosing ν such that the limit in (116) is equal to zero and the limit in (117) is equal to G ensures that the input distribution in (111) is first-order optimal in the sense of [34, Theorem 4]. Note that since the transmitter knows the main channel gain h, then $G = \sup\limits_{\gamma_h} \gamma_h$ [34]. The fact that the support of h is infinite (by assumption of Theorem 4) induces that $G = \infty$. The secrecy rate achieved by the above input distribution is given by:

$$R_{-2} = \mathop{\mathbf{E}}_{\gamma_h, x}\left[\log\left(1 + \gamma_h |x|^2\right)\right] - \mathop{\mathbf{E}}_{\gamma_g, x}\left[\log\left(1 + \gamma_g |x|^2\right)\right]. \tag{118}$$

As $P_{max} \to 0$, the first term in (118) is much larger than the second one as shown below:

$$\begin{aligned}
\lim_{P_{max}\to 0} \frac{\mathop{\mathbf{E}}\limits_{\gamma_g, x}\left[\log\left(1 + \gamma_g |x|^2\right)\right]}{\mathop{\mathbf{E}}\limits_{\gamma_h, x}\left[\log\left(1 + \gamma_h |x|^2\right)\right]} &= \lim_{P_{max}\to 0} \frac{\frac{\mathop{\mathbf{E}}\limits_{\gamma_g, x}\left[\log\left(1+\gamma_g |x|^2\right)\right]}{P_{max}}}{\frac{\mathop{\mathbf{E}}\limits_{\gamma_h, x}\left[\log\left(1+\gamma_h |x|^2\right)\right]}{P_{max}}} && (119)\\
&\leq \lim_{P_{max}\to 0} \frac{\frac{\log\left(1+\mathop{\mathbf{E}}\limits_{\gamma_g}\left[\gamma_g\right] P_{max}\right)}{P_{max}}}{\frac{\mathop{\mathbf{E}}\limits_{\gamma_h, x}\left[\log\left(1+\gamma_h |x|^2\right)\right]}{P_{max}}} && (120)\\
&= \frac{\mathop{\mathbf{E}}\limits_{\gamma_g}\left[\gamma_g\right]}{G} && (121)\\
&= 0, && (122)
\end{aligned}$$

where (120) is due to the Jensen's inequality and (121) follows because the input x is first-order optimal. Hence, R_{-2} is asymptotically equal to

$$R_{-2} \approx \mathop{\mathbf{E}}_{\gamma_h, x}\left[\log\left(1 + \gamma_h |x|^2\right)\right]. \tag{123}$$

The rate on the RHS of (123) is asymptotically equal to the capacity of the main channel and hence is the best rate one can achieve. To conclude the proof, we note that to set the input distribution (111), one only needs to know when the actual channel gain is above the threshold ν which is possible through a 1-bit feedback.

Appendix J
Proof of the statement in Remark 4

We prove the statement in Remark 4 via an example. Let us consider fading channels with PDF defined on $[0, a]$ by:

$$f_{\gamma_h}(x) = f_{\gamma_g}(x) = \frac{1}{a}, \tag{124}$$

where a is an arbitrary positive number. The capacity without secrecy over the main channel is given by [35]:

$$\begin{aligned}
C_{w.s}(P_{max}) &= \int_{\lambda}^{a} \log\left(\frac{\gamma_h}{\lambda}\right) \frac{1}{a}\, \mathrm{d}\gamma_h && (125)\\
&= -1 + \frac{\lambda}{a} + \log\left(\frac{a}{\lambda}\right), && (126)
\end{aligned}$$

where λ is the cut-off rate obtained by solving $\int_{\lambda}^{a}\left(\frac{1}{\lambda} - \frac{1}{\gamma_h}\right)\frac{1}{a}\ \mathrm{d}\gamma_h = P_{max}$. It can be verified that λ can be obtained explicitly as $\lambda = -\frac{a}{W\left(-e^{-1-a\,P_{max}}\right)}$, where $W(\cdot)$ is the principal branch of the LambertW function. Substituting the later expression of λ in (126), we obtain:

$$\begin{aligned}
C_{w.s}(P_{max}) &= -1 - \frac{1}{W\left(-e^{-1-a\,P_{max}}\right)} + \log\left(-W\left(-e^{-1-a\,P_{max}}\right)\right) && (127)\\
&= a\,P_{max} + \mathrm{o}(P_{max}) && (128)\\
&\approx a\,P_{max}, && (129)
\end{aligned}$$

where we have used the fact that $W\left(-e^{-1-a\,x}\right) = -1 + \sqrt{2\,a\,x} - \frac{2}{3}\,a\,x + \mathrm{o}(x)$ to obtain (128). Note that (129) is in full agreement with the framework in [34] since $G = a$ for the PDF's considered above. Next, we show that the secrecy capacity of this channel is at most asymptotically equal to $\frac{a}{2}\,P_{max}$ and is thus strictly smaller than the capacity without secrecy constraint. To that end, we upper-bound the secrecy capacity with perfect main CSI given in (7) as follows:

$$\begin{aligned}
\mathop{\mathbf{E}}_{\gamma_h, \gamma_g}\left[\left[\log\left(\frac{1 + \gamma_h\,P(h)}{1 + \gamma_g\,P(h)}\right)\right]^+\right] &\leq \mathop{\mathbf{E}}_{\gamma_h, \gamma_g}\left[\left[\log\left(\frac{1 + a\,P(h)}{1 + \gamma_g\,P(h)}\right)\right]^+\right] && (130)\\
&\leq \mathop{\mathbf{E}}_{\gamma_g}\left[\log\left(\frac{1 + a\mathop{\mathbf{E}}\limits_{\gamma_h}[P(h)]}{1 + \gamma_g \mathop{\mathbf{E}}\limits_{\gamma_h}[P(h)]}\right)\right], && (131)
\end{aligned}$$

where (131) follows from the Jensen's inequality since the function $x \mapsto \log\left(\frac{1+c\,x}{1+d\,x}\right)$ is concave for all $0 \leq d \leq c$. Since the RHS of (131) is increasing in $\underset{\gamma_h}{\mathbf{E}}[P(h)]$, then maximizing both sides of (131) with respect to the LTPC: $\underset{\gamma_h}{\mathbf{E}}[P(\boldsymbol{h})] \leq P_{max}$, we obtain:

$$R_{++} \leq \underset{\gamma_g}{\mathbf{E}}\left[\log\left(\frac{1+a\,P_{max}}{1+\gamma_g\,P_{max}}\right)\right] \quad (132)$$

$$\approx P_{max}\,\underset{\gamma_g}{\mathbf{E}}\left[a-\gamma_g\right] \quad (133)$$

$$= \frac{a}{2}\,P_{max}, \quad (134)$$

which we wanted to show.

Acknowledgment

The authors would like to thank the editor Dr. Shi Jin for volunteering his time to handle this paper and the anonymous reviewers for their valuable comments that have enhanced the technical quality and the lucidity of this paper. A special thanks is addressed to E. Koksal and O. Basciftci for correcting a previous version of the proof of Theorem 3, and H. Tembine for invoking the statistical dominance argument.

References

[1] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[2] M. Bloch, J. Barros, M. Rodrigues, and S. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

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

[4] M. Bloch and J. Laneman, "Information-spectrum methods for information-theoretic security," in *Proc. Information Theory and Applications Workshop (ITA'2009)*, San Diego, CA, USA, Feb. 2009, pp. 23–28.

[5] Z. Rezki, A. Khisti, and M.-S. Alouini, "On the ergodic secret message capacity of the wiretap channel with finite-rate feedback," in *Proc. 2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Cambridge, MA, USA, july 2012, pp. 239 –243.

[6] M. Bloch and J. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE J. Select. Areas Commun.*, vol. 31, no. 9, pp. 1840–1849, Sep. 2013.

[7] Z. Rezki, A. Khisti, and M.-S. Alouini, "On the ergodic secrecy capacity of the wiretap channel under imperfect main channel estimation," in *Proc. 2011 45th Asilomar Conference on Signals, Systems and Computers (Asilomar'2011)*, Pacific Grove, CA, USA, Nov 2011, pp. 952–957.

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

[9] X. Zhou and M. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[10] A. Mukherjee and A. Swindlehurst, "Ensuring secrecy in MIMO wiretap channels with imperfect CSIT: A beamforming approach," in *2010 IEEE International Conference on Communications (ICC'2010)*, Cape Town, South Africa, May 2010, pp. 1–5.

[11] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, Jul. 2012.

[12] A. Mukherjee and A. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Processing*, vol. 59, no. 1, pp. 351–361, Jan 2011.

[13] W. Shi and J. Ritcey, "Robust beamforming for MISO wiretap channel by optimizing the worst-case secrecy capacity," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR'2010)*, Monterey, CA, USA, 7-10 Nov. 2010, pp. 300–304.

[14] J. Taylor, M. Hempel, H. Sharif, S. Ma, and Y. Yang, "Impact of channel estimation errors on effectiveness of eigenvector-based jamming for physical layer security in wireless networks," in *the 16th IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD'2011)*, Kyoto, Japan, Jun. 2011, pp. 122–126.

[15] S.-C. Lin, T.-H. Chang, Y.-L. Liang, Y. Hong, and C.-Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy with artificial noise: The noise leakage problem," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 901–915, March 2011.

[16] S. Gerbracht, C. Scheunert, and E. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inform. Forensics and Security*, vol. 7, no. 2, pp. 704–716, 2012.

[17] J. Li and A. Petropulu, "Explicit solution of worst-case secrecy rate for MISO wiretap channels with spherical uncertainty," *IEEE Trans. Signal Processing*, vol. 60, no. 7, pp. 3892–3895, 2012.

[18] S. Ma, M. Hong, E. Song, X. Wang, and D. Sun, "Outage constrained robust secure transmission for MISO wiretap channels," *submitted for publication, available at http://arxiv.org/pdf/1310.7158.pdf*, 2013.

[19] C.-W. Huang, T.-H. Chang, X. Zhou, and Y.-W. Hong, "Two-way training for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2724–2738, May 2013.

[20] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[21] T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, Apr. 2007.

[22] D. Love, R. W. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, october 2008.

[23] V. Hassel, D. Gesbert, M.-S. Alouini, and G. Oien, "A threshold-based channel state feedback algorithm for modern cellular systems," *IEEE Trans. Wireless Commun.,*, vol. 6, no. 7, pp. 2422–2426, july 2007.

[24] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[25] A. Khisti and G. Wornell, "Secure transmission with multiple antennas. Part I: The MISOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[26] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," in *Proceedings of the IEEE International Symposium on Information Theory, Massachusetts, USA*, Aug. 1998, p. 42.

[27] Y. Abdallah, M. Latif, M. Youssef, A. Sultan, and H. El Gamal, "Keys through ARQ: Theory and practice," *IEEE Trans. Inform. Forensics and Security*, vol. 6, no. 3, pp. 737–751, Sep. 2011.

[28] I. Safaka, C. Fragouli, K. Argyraki, and S. Diggavi, "Creating shared secrets out of thin air," in *Proc. 11th ACM Workshop on Hot Topics in Networks, HotNets-XI*, Redmond, Seattle, WA, USA, Oct. 2012, pp. 73–78.

[29] R. Davidson and J.-Y. Duclos, "Testing for restricted stochastic dominance," *Econometric Reviews*, vol. 32, no. 1, pp. 84–125, 2013. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/07474938.2012.690332

[30] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Review*, vol. 44, no. 4, pp. 525–597, 2002.

[31] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, C. U. Press, Ed. Cambridge University Press, 2003.

[32] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, ser. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, July 2006.

[33] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[34] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[35] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1986 –1992, Nov. 1997.

Zouheir Rezki (S'01, M'08, SM'13) was born in Casablanca, Morocco. He received the Diplôme d'Ingénieur degree from the École Nationale de l'Industrie Minérale (ENIM), Rabat, Morocco, in 1994, the M.Eng. degree from École de Technologie Supérieure, Montreal, Québec, Canada, in 2003, and the Ph.D. degree from École Polytechnique, Montreal, Québec, in 2008, all in electrical engineering. From October 2008 to September 2009, he was a postdoctoral research fellow with Data Communications Group, Department of Electrical and Computer Engineering, University of British Columbia. He is now a Research Scientist at King Abdullah University of Science and Technology (KAUST), Thuwal, Mekkah Province, Saudi Arabia. His research interests include: performance limits of communication systems, physical-layer security, cognitive and sensor networks and low-complexity detection algorithms.

Ashish Khisti (S02, M09) is an Assistant Professor in the Electrical and Computer Engineering (ECE) Department and a Canada Research Chair (Tier II) in Network Information Theory at the University of Toronto, Toronto, Ontario, Canada. He received his BASc degree in Engineering Sciences from University of Toronto in 2002 and his S.M. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA in 2004 and 2008, respectively. He has been with the University of Toronto since 2009. His research interests span the areas of information theory, wireless physical layer security and streaming communication systems. During his graduate studies, Professor Khisti was a recipient of the NSERC postgraduate fellowship, Harold H. Hazen Teaching award and the Morris Joseph Levin Masterworks award. At the University of Toronto he is a recipient of the Ontario Early Researcher Award (2012) and a Hewlett-Packard IRP award (2011, 2012). He is an associate editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.

Mohamed-Slim Alouini (S'94, M'98, SM'03, F'09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.