On the secrecy capacity of the wiretap channel with imperfect main channel estimation

On the Secrecy Capacity of the Wiretap Channel with Imperfect Main Channel Estimation

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

***Abstract*—We study the secrecy capacity of fast fading channels under imperfect main channel (between the transmitter and the legitimate receiver) estimation at the transmitter. Lower and upper bounds on the ergodic secrecy capacity are derived for a class of independent identically distributed (i.i.d.) fading channels. The achievable rate follows from a standard wiretap code in which a simple on-off power control is employed along with a Gaussian input. The upper bound is obtained using an appropriate correlation scheme of the main and the eavesdropper channels, and is the best known upper bound so far. The upper and the lower bounds coincide with recently derived ones in case of perfect main CSI. Furthermore, the upper bound is tight in case of no main CSI, where the secrecy capacity is equal to zero. Asymptotic analysis at high and low signal-to-noise ratio (SNR) are also given. At high SNR, we show that the capacity is bounded by providing upper and lower bounds that depend on the channel estimation error. At low SNR, however, we prove that the secrecy capacity is asymptotically equal to the capacity of the main channel as if there were no secrecy constraint. Numerical results are provided for independent identically distributed (i.i.d.) Rayleigh fading channels.**

***Index Terms*—Secrecy capacity, imperfect channel estimation, noisy CSI, on-off signaling, fading channels, low SNR, high SNR.**

I. Introduction

The wiretap channel, in which a source communicates with a receiver through a discrete, memoryless channel (DMC) and a wire-tapper observes the output of this channel via another DMC, has been introduced by Wyner [1]. In this seminal work, it has been shown that if the capacity of the main channel (the channel between the transmitter and the legitimate receiver) is greater than the capacity of the wire-tapper, then there exists an encoding-decoding scheme such that reliable communication that keeps the messages completely secret against eavesdropping is possible (without the use of any encryption key). Leung-Yan-Cheong and Hellman have extended Wyner's work and characterized the secrecy-capacity and the achievable rate-equivocation region for the Gaussian wiretap channel with additive noise [2]. Later on, Csiszár generalized Wyner's wiretap channel by considering a non-degraded broadcast channel with confidential messages [3].

In terms of designing practical codes, secrecy-achieving codes have been proposed for some specific wiretap channels in e.g., [4], [5]. While the latter construction was based on low-density parity check (LDPC) codes, there has been recently an increasing effort toward explicit construction based on polar codes, e.g. [6]–[9], to cite only few. Constructions based on lattices are also proposed in e.g., [10], [11], whereas wiretap codes based on explicit extractors are presented in [12].

Motivated by these positive previous results, many other authors have recently addressed the impact of fading on secure communications. Intuitively, fading generally increases the randomness of the channel input and it is therefore not surprising that fading may help improve communication security. Indeed, it has been shown in, e.g., [13]–[15] that in a quasi-static fading channel and in contrast to the Gaussian channel, secure communication is possible even if the average signal-to-noise ratio (SNR) of the main channel is less than that of the wire-tapper (or one of the wire-tappers in a multiple eavesdroppers case as discussed in [16]). Moreover, if a high level of outage is to be tolerated, then the outage secrecy rate of the fading channel can even be higher than the secrecy capacity of the Gaussian wiretap channel for similar average SNR levels. The effect of fading on secure communication for single-antenna wiretap and broadcast channels has also been studied in [17]–[19] where the secrecy-capacity along with the optimal power allocation and/or rate-adaptation strategies at the source have been derived under different channel state information (CSI) assumptions. Most of the previous work, either assume perfect CSI at all terminals, or perfect main CSI at the transmitter (CSI-T).

The secrecy-capacity of multiple-antenna wiretap channels with fixed channel gains has been studied in [20]–[24]. The effect of multiple antennas in enhancing the security capability of a wireless link has also been addressed in [25], where the main channel is known to all terminals, but only the eavesdropper has access to its channel. The impact of multiple antennas on guaranteeing a certain level of security, in terms of low probability of intercept and low probability of detection constraints, has been addressed in [26]. The secret diversity-multiplexing tradeoff of a multiple-antenna wiretap channel has been investigated in [27]. Secure transmission schemes based on sending an artificial noise to enhance the eavesdropper equivocation are presented in, e.g., [28]–[32].

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 2011 IEEE Asilomar Conference on Signals, Systems, and Computers (Asilomar'2011), Pacific Grove, California, USA.

This work was also supported by the Qatar National Research Fund (a member of Qatar Foundation) under NPRP Grant NPRP 5-603-2-243. The statements made herein are solely the responsibility of the authors

Discussions on the effect of CSI estimation error on secrecy are also presented in [33]–[39].

The ergodic secrecy capacity of the wiretap channel is known when the main CSI is perfect at the transmitter, but for a sufficiently large coherence period [17]. To the best of our knowledge, the secrecy capacity in a fast fading scenario where the channel gains change from one symbol to the next, is only known when the transmitter is aware perfectly of both the main and the eavesdropper CSI [18]. It is still not known, if the eavesdropper CSI is not available, even with perfect main CSI at the transmitter. However, upper and lower bounds have been reported in [25]. Independently and concurrently with the conference version of this paper [40], an achievable rate has also been derived in [41] (see also [42]). In this paper, we study the secrecy capacity of fast fading channels under imperfect main channel estimation at the transmitter. More precisely, we assume that the main and the eavesdropper channels are independent identically distributed (i.i.d.), ergodic and stationary processes, with continuous and bounded probability density functions (PDF). Furthermore, we assume that the transmitter, in addition to the statistics of both channels, is also provided with an estimated value of the instantaneous main channel gain. The legitimate receiver is aware of its instantaneous channel gain along with the conditional received average SNR (defined formally later), whereas the eavesdropper's receiver, in addition to its instantaneous channel gain, is aware of what CSI the transmitter and the legitimate receiver have.

In the previous setting, we present upper and lower bounds on the secrecy capacity. The lower bound is obtained via a standard wiretap code [2], [3] (see also [15, Chap. 3]). The upper bound, which is our main contribution, follows by properly correlating the main and the eavesdropper channel gains. The upper bound depends on the main CSI-T estimation error and as shown numerically, it improves upon the upper bound corresponding to perfect main CSI-T. Furthermore, the upper and the lower bounds coincide with the ones derived in [25] in case of perfect main CSI. In addition, the upper bound is tight in case of no main CSI, where the secrecy capacity is equal to zero. Moreover, we provide asymptotic analysis in cases of perfect and no main CSI, together with results at high SNR and at low SNR. In the high-SNR regime, we show that the capacity is bounded. In the low-SNR regime, we find that the capacity is asymptotically equal to that of the main channel as if there were no eavesdropper, thus establishing the tightness of our bounds in this regime too.

We note that our model is different from Gopala et al.'s [17]. While [17] considers a block fading model where the coherence blocks are large enough to guarantee reliability in each of them, our focus is on a fast fading model. Recall that in fast fading channels, the coherence blocks are not necessarily large and thus even with perfect main channel state information (CSI) at the transmitter, reliability is not guaranteed in every coherence block. As a consequence, the achievability scheme in [17] is not applicable to our setting, even with perfect main CSI-T. It is evidently not applicable with a noisy CSI-T. Note that, in block fading channels and with perfect main CSI-T only (without the need of the eavesdropper CSI-T), the secrecy capacity is known [17]. In fast fading, however, provided that the eavesdropper CSI is not available at the transmitter, the secrecy capacity is still not known, even with perfect main CSI-T.

Contrasted with [41] which considers that CSI-T is a deterministic function of the exact CSI, our work deals with the setting where the CSI-T is a noisy version of the true CSI. Furthermore, Theorem 1 below provides a converse which is our main contribution, whereas there is no explicit converse in [41]. Additionally, there is no asymptotic analysis in [41], whereas our framework formally studies the high-SNR and the low-SNR regimes.

The organization of this paper is as follows. Section II introduces our system model, followed by our main result along with its proof in Section III. In Section IV, an asymptotic analysis is presented. Section V contains a summary of our results when applied to Rayleigh fading channels which we use in order to provide numerical results in Section VI. Finally, Section VII concludes the paper.

Notations: The expectation operation is denoted by $E[\cdot]$. 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. For a random variable x, $\boldsymbol{x}^n$ designates the vector $(x(1), \ldots, x(n))$. when there is no ambiguity, we find it also convenient to use x_i to designate $x(i)$. 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. The symbols $\gtrsim$ and $\lesssim$ are defined analogously. The functions $f_x(\cdot)$ and $F_x(\cdot)$ denote the probability density function (pdf) and the cumulative distribution function (cdf) of the random variable x. If x is a circularly symmetric Gaussian random variable with mean m and variance σ^2, then it is denoted as $x \sim \mathcal{CN}\left(m, \sigma^2\right)$.

II. System Model

We consider a discrete-time memoryless wiretap channel consisting of a transmitter, a legitimate receiver and an eavesdropper. Each terminal is equipped with a single antenna. The outputs at both the legitimate destination and the eavesdropper, at time period i, $i = 1, \ldots, n$, are expressed, respectively by:

$$\begin{cases} y(i) = h(i)\, x(i) + v_y(i) \\ z(i) = g(i)\, x(i) + v_z(i), \end{cases} \tag{1}$$

where $x(i) \in \mathbb{C}$ is the transmitted signal, 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, respectively; and $v_y(i) \in \mathbb{C}$, $v_z(i) \in \mathbb{C}$ are zero-mean, unit-variance circularly symmetric white Gaussian noises. The channel gains h and g are assumed to be i.i.d., ergodic and stationary with bounded and continuous PDF's. While the transmitter is aware of both h and g statistics, it is not aware of the channel realization g(i) and is only provided with a noisy version of $h(i)$, say $\hat{h}(i) \sim \mathcal{CN}(0, 1)$, such that the main channel estimation model can be written as:

$$h(i) = \sqrt{1-\alpha}\ \hat{h}(i) + \sqrt{\alpha}\ \tilde{h}(i), \tag{2}$$

where $\tilde{h}(i) \sim \mathcal{CN}(0, 1)$ is the estimation error and α is the error variance ($\alpha \in (0, 1)$). We assume that $\hat{h}(i)$ and $\tilde{h}(i)$ are

uncorrelated and thus independent. On the other hand, at the receiver sides, we assume that the legitimate receiver is aware of both its instantaneous channel gain $h(i)$ and its conditional received average SNR $\bar{h}(i)$ given by $\bar{h}(i) = |h(i)|^2\ \omega(i)$, where $\omega(i) = E\left[|x(i)|^2 \mid \hat{\boldsymbol{h}}^i\right]$ and the last expectation is over the conditional input distribution; whereas the eavesdropper's receiver, in addition to its instantaneous channel realizations $g(i)$, is aware of $\hat{h}(i)$, $h(i)$ and $\bar{h}(i)$. Our motivation to reveal $\hat{h}(i)$ to the eavesdropper is driven by the fact that the latter may be able to track the feedback link between the legitimate receiver and the transmitter, and thus it may retrieve $\hat{h}(i)$ exactly like at the source. As pointed out in [41], revealing $h(i)$ to the eavesdropper also prevents the legitimate terminals to use the CSI as a source of randomness for key generation. Finally, the channel input is constrained according to an average power constraint:

$$\frac{1}{n}\sum_{i=1}^{n} E\left[|x_i|^2\right] \leq P_{avg}.$$

We are interested in message transmission secrecy capacity of such a channel when $n \to \infty$. The level of uncertainty about the message w at the eavesdropper is measured by the (normalized) leakage of information that the eavesdropper gets about the message by observing its channel output, i.e., $\frac{1}{n} I\left(w; \boldsymbol{z}^n, \boldsymbol{g}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \mid \hat{\boldsymbol{h}}^n\right)$, where $\frac{1}{n} I\left(w; \boldsymbol{z}^n, \boldsymbol{g}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \mid \hat{\boldsymbol{h}}^n\right)$ denotes the mutual information between w and $\left(\boldsymbol{z}^n, \boldsymbol{g}^n, \hat{\boldsymbol{h}}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n\right)$. The eavesdropper is ignorant about the message if:

$$\lim_{n\to\infty} \frac{1}{n} I\left(w; \boldsymbol{z}^n, \boldsymbol{g}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \mid \hat{\boldsymbol{h}}^n\right) = 0. \tag{3}$$

A rate R is an achievable secrecy rate if there exists a sequence of $\left(n, 2^{nR}\right)$ codes, for which 2^{nR} represents the number of messages to be sent to the destination, such that (3) holds true and $\lim_{n\to\infty} P_e = 0$, where P_e is the average error probability defined by:

$$P_e = \frac{1}{2^{nR}} \sum_{w_0=1}^{2^{nR}} \Pr\{w \neq \hat{w} \mid w = w_0\}, \tag{4}$$

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: $C_s := \sup_{R\in\mathcal{R}_s} R$, where $\mathcal{R}_s$ is the set of achievable secrecy rates.

Remark 1: While our focus in this paper is on a weak secrecy constraint since the secrecy constraint in (3) is normalized by n, there exists a stronger secrecy measure that defines the secrecy in term of a mutual information, i.e., $\lim_{n\to\infty} I\left(w; \boldsymbol{z}^n, \boldsymbol{g}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \mid \hat{\boldsymbol{h}}^n\right) = 0$. For references considering the latter as well as other stronger notions of secrecy, please see, e.g., [43]–[45].

III. Ergodic Capacity

In this section, our main result is presented in Theorem 1, followed by the proof.

Theorem 1: The secrecy capacity of the discrete-time memoryless channel described by (1), under imperfect main channel estimation (2), is bounded as follows:

$$R_- \leq C_s \leq R_+, \tag{5}$$

where R_- and R_+ are given by:

$$R_- = \underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\log\left(\frac{1+P_0(\tau)|h|^2}{1+P_0(\tau)|g|^2}\right)\right] \tag{6}$$

$$R_+ = \max_{P(\hat{h})} \underset{\hat{h},\tilde{h}}{E}\left[\log\left(\frac{1+P(\hat{h})|\sqrt{1-\alpha}\,\hat{h}+\sqrt{\alpha}\,\tilde{h}|^2}{1+P(\hat{h})|\tilde{h}|^2}\right)\right]^+, \tag{7}$$

where $P_0(\tau) = \frac{P_{avg}}{1-F_{|\hat{h}|^2}(\tau)}$ and where τ can be optimized to maximize R_-.

Proof:

- Achievable rate:

To prove that R_- is achievable and following [46, Proposition3], we consider a new wiretap channel in which the input is amplified by $\sqrt{P\left(\hat{h}\right)}$, where $P(\cdot)$ is a time-invariant deterministic function that satisfies the power constraint. This amplifier may be regarded as part of the channel with input t and outputs y' and z' such that: $y'(i) = h'(i)\ t(i) + v_y(i)$ and $z'(i) = g'(i)\ t(i) + v_z(i)$, where we define $h'(i) = \sqrt{P\left(\hat{h}(i)\right)}\, h(i)$ and $g'(i) = \sqrt{P\left(\hat{h}(i)\right)}\, g(i)$, for convenience. This new channel has no CSI-T and input constraint $E\left[|t_i|^2\right] = 1$. By our CSI assumptions (cf. Section II), the legitimate receiver is aware of both the instantaneous channel gain $h(i)$ and its conditional received average SNR, say $\bar{h}'(i)$, which can be computed as follows:

$$\bar{h}'(i) = E\left[\frac{|E[y'(i) \mid t(i), h'(i)]|^2}{\mathrm{var}(y'(i) \mid t(i), h'(i))} \mid h'(i)\right] \tag{8}$$

$$= |h'(i)|^2 \tag{9}$$

$$= P\left(\hat{h}(i)\right)|h(i)|^2. \tag{10}$$

Knowing both $h(i)$ and $\bar{h}'(i)$, the legitimate receiver can construct $h'(i)$ and thus has perfect main CSI. Applying the result in [3] to the new channel and choosing $\boldsymbol{t}^n$ as an i.i.d. sequence with $t(i) \sim \mathcal{CN}(0,1)$, we get that

$$R_s = I(t; y', h') - I(t; z', g') \tag{11}$$

$$= I(t; y' \mid h') - I(t; z' \mid g') \tag{12}$$

$$= \underset{h,\hat{h},g}{E}\left[\log\left(1+P\left(\hat{h}\right)|h|^2\right) - \log\left(1+P\left(\hat{h}\right)|g|^2\right)\right], \tag{13}$$

is achievable. The rate in (13) can then be maximized over all power functions $P(\cdot)$ such that $E\left[P\left(\hat{h}\right)\right] \leq P_{avg}$. For simplicity, an on-off power scheme is adopted, i.e.,

$$P(\hat{h}) = \begin{cases} P_0(\tau) & |\hat{h}|^2 \geq \tau \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

which when applied to (13) yields (6). To complete the proof, the secrecy rate R_- is maximized over all positive τ values, and the optimum τ_0 is obtained by differentiating (6) with respect to τ. That is, τ_0 is a solution of

$$\underset{|h|^2,|\hat{h}|^2\geq\tau}{E}\left[\frac{P_0'(\tau_0)|h|^2}{1+P_0(\tau_0)|h|^2}\right] - \underset{|g|^2}{E}\left[\frac{P_0'(\tau_0)|g|^2}{1+P_0(\tau_0)|g|^2}\right]\left(1-F_{|\hat{h}|^2}(\tau_0)\right)$$

$$-f_{|\hat{h}|^2}(\tau_0)\left(\underset{|h|^2|\hat{h}|^2}{E}\left[\log\left(1+P_0(\tau_0)|h|^2\right)\mid|\hat{h}|^2=\tau_0\right]\right.$$
$$\left.-\underset{|g|^2}{E}\left[\log\left(1+P_0(\tau_0)|g|^2\right)\right]\right)=0, \quad (15)$$

where $P_0'(\tau)$ is the derivative of $P_0(\tau)$ with respect to τ.

- Upper bound

We recall that in our setting, at time instant i, the transmitter has CSI $\hat{h}_i$, whereas the legitimate receiver knows $\left(h_i, \bar{h}_i\right)$ and the eavesdropper's receiver knows $\left(g_i, \hat{h}_i, h_i, \bar{h}_i\right)$. Next, we prove the converse by following three main steps:

- Step 1: We consider an enhanced wiretap channel with higher secrecy capacity than the original channel.
- Step 2: We clearly specify the distributions upon which the secrecy and reliability conditions for the enhanced channel depend.
- Step 3: We show that substituting $\boldsymbol{g}^n$ by $\tilde{\boldsymbol{h}}^n$ in the enhanced channel does not change the secrecy capacity.
- Step 4: We construct a new wiretap channel where the above distributions are the same and where $\boldsymbol{g}^n$ is substituted by $\tilde{\boldsymbol{h}}^n$ (thus the secrecy capacity of the newly constructed wiretap channel is equal to that of the enhanced channel), but where the noises of the main channel and the eavesdropper's channel are correlated such that the new wiretap channel is degraded.
- Step 5: We upper-bound the secrecy capacity of the new wiretap channel, thereby proving the converse.

Step 1:
The enhanced channel is a fading wiretap channel similar to the one given by (1), where the main CSI-T is given by (2), but where the eavesdropper is not aware of $\boldsymbol{h}^n$ and $\bar{\boldsymbol{h}}^n$. That is, the secrecy constraint of the enhanced channel is defined by:

$$\lim_{n\to\infty}\frac{1}{n}I\left(w;\boldsymbol{z}^n,\boldsymbol{g}^n\mid\hat{\boldsymbol{h}}^n\right)=0. \quad (16)$$

Clearly, the secrecy capacity of the enhanced channel is at least equal to that of the original one.

Step 2:
Suppose that a rate R is achievable on the enhanced channel. Then, there exists a sequence of $\left(n, 2^{nR}, \delta_n\right)$ codes, such that:

$$\begin{cases}\frac{1}{n}H\left(w\mid\hat{\boldsymbol{h}}^n\right)-\frac{1}{n}H\left(w\mid\boldsymbol{z}^n,\boldsymbol{g}^n,\hat{\boldsymbol{h}}^n\right)\le\delta_n\\ P_e\le\delta_n,\end{cases} \quad (17)$$

for a sequence δ_n, with $\delta_n\to 0$ as $n\to\infty$. The secrecy condition in (17) depends on the joint distribution $p_{\boldsymbol{z}^n,\boldsymbol{g}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$ which in regard of the fact that the eavesdropper's channel is not known at the transmitter and thus $\boldsymbol{g}^n$ is independent of $\left(w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n\right)$, can be written as:

$$p_{\boldsymbol{z}^n,\boldsymbol{g}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}=p_w p_{\boldsymbol{x}^n|w,\hat{\boldsymbol{h}}^n}p_{\hat{\boldsymbol{h}}^n}p_{\boldsymbol{g}^n}p_{\boldsymbol{z}^n|\boldsymbol{x}^n,\boldsymbol{g}^n}. \quad (18)$$

Similarly, the reliability condition in (17) depends on the joint distribution $p_{\boldsymbol{y}^n,\boldsymbol{h}^n,\hat{\boldsymbol{h}}^n,\bar{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$ which itself decomposes into:

$$p_{\boldsymbol{y}^n,\boldsymbol{h}^n,\hat{\boldsymbol{h}}^n,\bar{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}=p_w p_{\boldsymbol{x}^n|w,\hat{\boldsymbol{h}}^n}p_{\hat{\boldsymbol{h}}^n}p_{\boldsymbol{h}^n|\hat{\boldsymbol{h}}^n}p_{\boldsymbol{y}^n|\boldsymbol{x}^n,\boldsymbol{h}^n}. \quad (19)$$

Any new wiretap channel that preserves the distributions (18) and (19) satisfies the secrecy and the reliability constraints in (17) and thus has equal secrecy capacity as the enhanced wiretap channel.

Step 3:
For the enhanced wiretap channel, we note that substituting $\boldsymbol{g}^n$ by $\tilde{\boldsymbol{h}}^n$ preserves the marginal (18) and (19). This is formalized by Lemma 1.

Lemma 1: Consider a wiretap channel defined by:

$$\begin{cases}y(i)=h(i)\,x(i)+v_y(i)\\ z(i)=\tilde{h}(i)\,x(i)+v_z(i),\end{cases} \quad (20)$$

with similar CSI as the wiretap channel (1), error probability (4) and secrecy constraint defined by $\lim_{n\to\infty}\frac{1}{n}I\left(w;\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n\mid\hat{\boldsymbol{h}}^n\right)=0$. Then the secrecy capacity of this channel and that of the enhanced channel described above are equal.

Proof: To prove Lemma 1, we only need to verify that $p_{\boldsymbol{z}^n,\boldsymbol{g}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}=p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$ since $p_{\boldsymbol{y}^n,\boldsymbol{h}^n,\hat{\boldsymbol{h}}^n,\bar{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$ is the same for both channels.

$$\begin{aligned}p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}&=p_w p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n|w} &(21)\\ &=p_w p_{\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n|w}p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n|w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n} &(22)\\ &=p_w p_{\hat{\boldsymbol{h}}^n}p_{\boldsymbol{x}^n|w,\hat{\boldsymbol{h}}^n}p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n|w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n} &(23)\\ &=p_w p_{\hat{\boldsymbol{h}}^n}p_{\boldsymbol{x}^n|w,\hat{\boldsymbol{h}}^n}p_{\tilde{\boldsymbol{h}}^n|w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n}p_{\boldsymbol{z}^n|w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,\tilde{\boldsymbol{h}}^n} &(24)\\ &=p_w p_{\hat{\boldsymbol{h}}^n}p_{\boldsymbol{x}^n|w,\hat{\boldsymbol{h}}^n}p_{\tilde{\boldsymbol{h}}^n}p_{\boldsymbol{z}^n|\boldsymbol{x}^n,\tilde{\boldsymbol{h}}^n}, &(25)\end{aligned}$$

where (23) follows because w and $\hat{\boldsymbol{h}}^n$ are independent and (25) holds true due to our CSI assumption and the fact that $\tilde{\boldsymbol{h}}^n$ and $\hat{\boldsymbol{h}}^n$ are independent, thus $\tilde{\boldsymbol{h}}^n$ and $\left(w,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n\right)$ are independent. Comparing the right hand side (RHS) of (18) and (25), we note that $p_{\boldsymbol{g}^n}=p_{\tilde{\boldsymbol{h}}^n}$ since both g and $\tilde{h}$ are $\mathcal{CN}(0,1)$. Moreover, since $z\mid x,g$ follows a $\mathcal{CN}(xg,1)$ and since $z\mid x,\tilde{h}$ follows a $\mathcal{CN}\left(x\tilde{h},1\right)$, then $p_{\boldsymbol{z}^n|\boldsymbol{x}^n,\boldsymbol{g}^n}$ and $p_{\boldsymbol{z}^n|\boldsymbol{x}^n,\tilde{\boldsymbol{h}}^n}$ are statistically equivalent. Hence, the RHS of (18) and (25) are equal and so are $p_{\boldsymbol{z}^n,\boldsymbol{g}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$ and $p_{\boldsymbol{z}^n,\tilde{\boldsymbol{h}}^n,\hat{\boldsymbol{h}}^n,\boldsymbol{x}^n,w}$. ■

Next, we construct a new wiretap channel which satisfies (18) and (19) and where $\boldsymbol{g}^n$ is substituted by $\tilde{\boldsymbol{h}}^n$.

Step 4:
Following [18], the new wiretap channel is a fading wiretap channel similar to the enhanced one (also similar to the original one, but with secrecy constraint defined by (16)) where $\boldsymbol{g}^n$ is substituted by $\tilde{\boldsymbol{h}}^n$ and where the noises v_y and v_z are correlated. More specifically, the new channel is defined as follows:

$$\begin{cases}y(i)=h(i)\,x(i)+v_y(i)\\ z(i)=\frac{\tilde{h}(i)h^*(i)}{|h(i)|^2}\left(h(i)\,x(i)+v_y(i)\right)+v_z'(i),\end{cases}\quad\text{if }\left(h(i),\tilde{h}(i)\right)\in\mathcal{A}_i, \quad (26)$$

and

$$\begin{cases}y(i)=\frac{h(i)\tilde{h}^*(i)}{|\tilde{h}(i)|^2}\left(\tilde{h}(i)\,x(i)+v_z(i)\right)+v_y'(i)\\ z(i)=\tilde{h}(i)\,x(i)+v_z(i),\end{cases}\quad\text{if }\left(h(i),\tilde{h}(i)\right)\notin\mathcal{A}_i, \quad (27)$$

where $\mathcal{A}_i=\left\{\left(h(i),\tilde{h}(i)\right):|h(i)|>|\tilde{h}(i)|\right\}$ and $v_z'(i)\sim\mathcal{CN}\left(0,1-\frac{|\tilde{h}(i)|^2}{|h(i)|^2}\right)$ and $v_y'(i)\sim\mathcal{CN}\left(0,1-\frac{|h(i)|^2}{|\tilde{h}(i)|^2}\right)$. Note that the channel defined by (26)-(27) satisfies both (18) and (19) and hence any secrecy rate achieved on the enhanced channel is also achieved on the new one. Furthermore, the new channel is physically degraded, i.e., $x(i)\to y(i)\to z(i)$ if $\left(h(i),\tilde{h}(i)\right)\in\mathcal{A}_i$

and $x(i) \to z(i) \to y(i)$ if $\left(h(i), \tilde{h}(i)\right) \notin \mathcal{A}_i$.

Step 5:

We now upper bound the secrecy rate of the new channel. as follows:

$$
\begin{aligned}
nR &= H\left(w \mid \hat{\boldsymbol{h}}^n\right) && (28)\\
&= H\left(w \mid \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n\right) + I\left(w; \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n \mid \hat{\boldsymbol{h}}^n\right) && (29)\\
&\leq H\left(w \mid \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n\right) + n\delta_n && (30)\\
&\leq I\left(w; \boldsymbol{y}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \,\middle|\, \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n\right) + n\left(\delta_n + \delta_n^{'}\right) && (31)\\
&\leq I\left(\boldsymbol{x}^n; \boldsymbol{y}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n \,\middle|\, \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n\right) + n\,\epsilon_n && (32)\\
&= I\left(\boldsymbol{x}^n; \boldsymbol{y}^n \,\middle|\, \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n\right) + n\,\epsilon_n && (33)\\
&= \sum_{i=1}^{n} \Big\{ h\left(y_i \,\middle|\, \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n, \boldsymbol{y}^{i-1}\right)\\
&\quad - h\left(y_i \,\middle|\, \boldsymbol{z}^n, \tilde{\boldsymbol{h}}^n, \hat{\boldsymbol{h}}^n, \boldsymbol{h}^n, \bar{\boldsymbol{h}}^n, \boldsymbol{y}^{i-1}, \boldsymbol{x}^n\right)\Big\} + n\,\epsilon_n && (34)\\
&\leq \sum_{i=1}^{n} \Big\{ h\left(y_i \,\middle|\, z_i, \tilde{h}_i, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right)\\
&\quad - h\left(y_i \,\middle|\, z_i, \tilde{h}_i, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i, x_i\right)\Big\} + n\,\epsilon_n && (35)\\
&= \sum_{i=1}^{n} I\left(x_i; y_i \,\middle|\, z_i, \tilde{h}_i, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) + n\,\epsilon_n && (36)\\
&= \sum_{i=1}^{n} \left[I\left(x_i; y_i \,\middle|\, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) - I\left(x_i; z_i \,\middle|\, \tilde{h}_i, \hat{\boldsymbol{h}}^i\right)\right]^+ + n\,\epsilon_n && (37)\\
&= \sum_{i=1}^{n} \left[h\left(y_i \,\middle|\, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) - h\left(z_i \,\middle|\, \tilde{h}_i, \hat{\boldsymbol{h}}^i\right)\right]^+ + n\,\epsilon_n && (38)\\
&\leq \sum_{i=1}^{n} \max_{p(\boldsymbol{x}(i)|\hat{\boldsymbol{h}}^i)} \left[h\left(y_i \,\middle|\, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) - h\left(z_i \,\middle|\, \tilde{h}_i, \hat{\boldsymbol{h}}^i\right)\right]^+ + n\,\epsilon_n && (39)\\
&= \sum_{i=1}^{n} E\left[\left[\log\left(\frac{1+|h_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}{1+|\tilde{h}_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}\right)\right]^+\right] + n\,\epsilon_n, && (40)
\end{aligned}
$$

where (30) follows from the secrecy constraint (17) and from substituting $\boldsymbol{g}^n$ by $\tilde{\boldsymbol{h}}^n$, and (31) follows from Fano's inequality, with $\delta_n^{'} \to 0$ as $n \to \infty$. For convenience, we have defined $\epsilon_n = \delta_n + \delta_n^{'}$ after (31). Inequality (32) is true because given $\hat{\boldsymbol{h}}^n$, $w \to \boldsymbol{x}^n \to \left(\boldsymbol{y}^n, \boldsymbol{h}^n, \hat{\boldsymbol{h}}^n\right)$ forms a Markov chain, whereas (33) holds since given $\hat{\boldsymbol{h}}^n$, $\boldsymbol{x}^n$ is independent of $\left(\boldsymbol{h}^n, \bar{\boldsymbol{h}}^n\right)$. In order to justify (37), we used the fact that the new channel is physically degraded and thus $I\left(x_i; y_i \,\middle|\, z_i, \tilde{h}_i, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) = \left[I\left(x_i; y_i \,\middle|\, \hat{\boldsymbol{h}}^i, h_i, \bar{h}_i\right) - I\left(x_i; z_i \,\middle|\, \tilde{h}_i, \hat{\boldsymbol{h}}^i\right)\right]^+$. Since given h_i and $\bar{h}_i$, the variance of y_i is equal to $1 + |h_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)$, where $P_i\left(\hat{\boldsymbol{h}}^i\right) = E\left[|x_i|^2 \,\middle|\, \hat{\boldsymbol{h}}^i\right]$, since given $\left(\tilde{h}_i, \hat{\boldsymbol{h}}^i\right)$, the variance of z_i is equal to $1 + |\tilde{h}_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)$, and since the input distribution that maximizes (39) is Gaussian [22], [23], [25], [47], then (40) holds true. Hence, it only remains to prove that the above upper bound is maximized by a power allocation $P_i\left(\hat{\boldsymbol{h}}^i\right) = \mu\left(\hat{h}_i\right)$, a time-invariant function of $\hat{h}_i$ only. To do this, we have:

$$
\begin{aligned}
&E\left[\left[\log\left(\frac{1+|h_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}{1+|\tilde{h}_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}\right)\right]^+\right]\\
&= E\left[E\left[\left[\log\left(\frac{1+|h_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}{1+|\tilde{h}_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right)}\right)\right]^+ \,\middle|\, |h_i|^2, |\tilde{h}_i|^2, \hat{h}_i\right]\right] && (41)\\
&\leq E\left[\left[\log\left(\frac{1+E\left[|h_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right) \,\middle|\, |h_i|^2, |\tilde{h}_i|^2, \hat{h}_i\right]}{1+E\left[|\tilde{h}_i|^2\, P_i\left(\hat{\boldsymbol{h}}^i\right) \,\middle|\, |h_i|^2, |\tilde{h}_i|^2, \hat{h}_i\right]}\right)\right]^+\right] && (42)\\
&= E\left[\left[\log\left(\frac{1+|h_i|^2\, E\left[P_i\left(\hat{\boldsymbol{h}}^i\right) \,\middle|\, \hat{h}_i\right]}{1+|\tilde{h}_i|^2\, E\left[P_i\left(\hat{\boldsymbol{h}}^i\right) \,\middle|\, \hat{h}_i\right]}\right)\right]^+\right] && (43)\\
&= E\left[\left[\log\left(\frac{1+|h_i|^2\, \mu_i\left(\hat{h}_i\right)}{1+|\tilde{h}_i|^2\, \mu_i\left(\hat{h}_i\right)}\right)\right]^+\right] && (44)\\
&= E\left[\left[\log\left(\frac{1+|h|^2\, \mu_i\left(\hat{h}\right)}{1+|\tilde{h}|^2\, \mu_i\left(\hat{h}\right)}\right)\right]^+\right] && (45)
\end{aligned}
$$

where (42) follows from Jensen's inequality since the function $x \mapsto \left[\log\left(\frac{1+a\,x}{1+b\,x}\right)\right]^+$ is concave for any positive a and b; where (43) follows because conditioned on $\hat{h}_i$, $\hat{\boldsymbol{h}}^i$ is independent of $|h_i|^2$ and $|\tilde{h}_i|^2$ due to the fact that the fading process $\{h_i\}$ is i.i.d.; where we have defined $\mu_i\left(\hat{h}_i\right)$ in (44) as $\mu_i\left(\hat{h}_i\right) = E\left[P_i\left(\hat{\boldsymbol{h}}^i\right) \,\middle|\, \hat{h}_i\right]$. Since the fading processes $\{h_i\}$, $\left\{\hat{h}_i\right\}$ and $\left\{\tilde{h}_i\right\}$ are ergodic and stationary, then they have a stationary first-order distribution and thus the expectation in (44) does not depend on their time index i which yields (45). Combining (40) and (45), we obtain:

$$
\begin{aligned}
R_e &\leq \frac{1}{n}\sum_{i=1}^{n} E\left[\left[\log\left(\frac{1+|h|^2\, \mu_i\left(\hat{h}\right)}{1+|\tilde{h}|^2\, \mu_i\left(\hat{h}\right)}\right)\right]^+\right] + \delta_n && (46)\\
&\leq E\left[\left[\log\left(\frac{1+|h|^2\, \frac{1}{n}\sum_{i=1}^{n} \mu_i\left(\hat{h}\right)}{1+|\tilde{h}|^2\, \frac{1}{n}\sum_{i=1}^{n} \mu_i\left(\hat{h}\right)}\right)\right]^+\right] + \delta_n && (47)\\
&= E\left[\left[\log\left(\frac{1+|h|^2\, \mu\left(\hat{h}\right)}{1+|\tilde{h}|^2\, \mu\left(\hat{h}\right)}\right)\right]^+\right] + \delta_n && (48)
\end{aligned}
$$

where (47) follows again by Jensen's inequality and where $\mu\left(\hat{h}\right)$ in (48) is defined as $\mu\left(\hat{h}\right) = \frac{1}{n}\sum_{i=1}^{n} \mu_i\left(\hat{h}\right)$. The above upper bound is tight if $\mu_i\left(\hat{h}\right)$ is independent of i. Letting $n \to \infty$ and maximizing over all power policies $\left\{\mu\left(\hat{h}\right)\right\}$ that satisfy the power constraint, we establish that

$$
R_e \leq \max_{E[\mu(\hat{h})] \leq P_{avg}} \underset{h,g,\hat{h}}{E}\left[\log\left(\frac{1+|h|^2\, \mu\left(\hat{h}\right)}{1+|\tilde{h}|^2\, \mu\left(\hat{h}\right)}\right)\right]^+. \qquad (49)
$$

That is, the following upper bound holds true:

$$
C_s \leq \max_{E[P(\hat{h})] \leq P_{avg}} \underset{h,g,\hat{h}}{E}\left[\log\left(\frac{1+P\left(\hat{h}\right)|h|^2}{1+P\left(\hat{h}\right)|\tilde{h}|^2}\right)\right]^+. \qquad (50)
$$

Note that the obtained upper bound, being an upper bound on the enhanced channel, is also an upper bound on the original channel which yields the result in (7). ■

It is worth mentioning that the upper bound has the following interpretation. In order to increase the information leakage, the eavesdropper "sticks" to the component of the

main channel that is unknown to the transmitter. Note also that the lower and the upper bounds on the secrecy capacity in (5) coincide with those derived in [25, Theorem 4], for i.i.d. fading channels under perfect main CSI. Although the lower and the upper bounds do not generally coincide, they provide, to the best of our knowledge, the best available characterization of the secrecy capacity over i.i.d. fading channels that accounts for imperfect main channel estimation at the transmitter. Note also that since R_+ in (7) is concave in $P\left(\hat{h}\right)$, the maximum can be found by deriving the optimum power profile $P(\hat{h})$ using the Lagrange approach. Indeed, the corresponding Lagrangian can be written as:

$$\mathcal{L}\left(P\left(\hat{h}\right),\lambda\right) = \underset{\hat{h}}{E}\left[\underset{\tilde{h}\,|\,\hat{h}}{E}\left[\left[\log\left(\frac{1+P(\hat{h})|\sqrt{1-\alpha}\;\hat{h}+\sqrt{\alpha}\;\tilde{h}|^2}{1+P(\hat{h})|\tilde{h}|^2}\right)\right]^+\middle|\hat{h}\right]\right] - \lambda\left(\underset{\hat{h}}{E}\left[P(\hat{h})\right]-P_{avg}\right), \tag{51}$$

where $\lambda \geq 0$ is the Lagrange multiplier corresponding to the average power constraint. Differentiating $\mathcal{L}\left(P\left(\hat{h}\right),\lambda\right)$ with respect to $P\left(\hat{h}\right)$ yields the following necessary and sufficient condition for optimality:

$$\frac{\partial}{\partial P\left(\hat{h}\right)}\left(\underset{\tilde{h}\,|\,\hat{h}}{E}\left[\left[\log\left(\frac{1+P(\hat{h})|\sqrt{1-\alpha}\;\hat{h}+\sqrt{\alpha}\;\tilde{h}|^2}{1+P(\hat{h})|\tilde{h}|^2}\right)\right]^+\middle|\hat{h}\right]\right)-\lambda=0. \tag{52}$$

For convenience, let $\hat{h}=\hat{\rho}e^{i\hat{\theta}}$ with $\hat{\rho}\in[0,\infty)$ and $\hat{\theta}\in[-\pi,\pi]$, and let us define the region $D_{\hat{h}}$ by:

$$D_{\hat{h}} = \left\{\tilde{h}=\tilde{\rho}e^{i\tilde{\theta}} \;\middle|\; \tilde{\rho}\leq\frac{\hat{\rho}}{\rho_0\left(\hat{\theta}-\tilde{\theta}\right)}\right\}, \tag{53}$$

where $\tilde{\rho}\in[0,\infty)$ and $\tilde{\theta}\in[-\pi,\pi]$ and $\rho_0(\cdot)$ is the function defined on $[-2\pi,2\pi]$ by:

$$\rho_0(t)=\frac{\sqrt{(1-\alpha)\left(\alpha\cos(t)^2-\alpha+1\right)}-\sqrt{\alpha(1-\alpha)}\cos(t)}{1-\alpha}. \tag{54}$$

Then, the optimal power profile is the solution of the following optimality condition:

$$\underset{\tilde{h}\in D_{\hat{h}}}{E}\left[\frac{|\sqrt{1-\alpha}\;\hat{h}+\sqrt{\alpha}\;\tilde{h}|^2}{1+P(\hat{h})|\sqrt{1-\alpha}\;\hat{h}+\sqrt{\alpha}\;\tilde{h}|^2}-\frac{|\tilde{h}|^2}{1+P(\hat{h})|\tilde{h}|^2}\right]-\lambda=0. \tag{55}$$

If for a particular value of $\hat{h}$, there is no positive solution $P\left(\hat{h}\right)$ for (55), then the instantaneous power is set to zero, i.e., $P\left(\hat{h}\right)=0$.

Remark 2: The achievable rate in Theorem 1 can be immediately improved by optimizing the rate R_s in (13) over all power policies that satisfy the power constraint. Indeed, departing from (13), one expect a better rate by solving:

$$\max_{E\left[P\left(\hat{h}\right)\right]\leq P_{avg}}\underset{h,\hat{h},g}{E}\left[\log\left(1+P\left(\hat{h}\right)\;|h|^2\right)-\log\left(1+P\left(\hat{h}\right)\;|g|^2\right)\right]. \tag{56}$$

Note that the objective function in (56) is not convex. Nevertheless, developing the Lagrangian exactly as in (51), it is possible to obtain the following necessary optimality condition via the Karush–Kuhn–Tucker (KKT) condition:

$$\underset{|h|^2|\hat{h},|g|^2}{E}\left[\frac{|h|^2-|g|^2}{\left(1+P(\hat{h})|h|^2\right)\left(1+P(\hat{h})|g|^2\right)}\right]=\lambda, \tag{57}$$

where λ is the Lagrange multiplier corresponding to the average power constraint. Define the function $f_{\hat{h}}(\cdot)$ as the LHS of (57), i.e., $f_{\hat{h}}(p)=\underset{|h|^2|\hat{h},|g|^2}{E}\left[\frac{|h|^2-|g|^2}{(1+P(\hat{h})|h|^2)(1+P(\hat{h})|g|^2)}\right]$. Then, following similar lines as [41, Lemma 5], one can show that if there exists $\hat{h}_0$, such that $E\left[|h|^2-|g|^2\mid\hat{h}_0\right]>0$, i.e., such that $(1-\alpha)\left(|\hat{h}_0|^2-1\right)>0$, then using the entire available power is optimal, and the power $p(\hat{h})$ defined by:

$$p(\hat{h})=\begin{cases}f_{\hat{h}}^{-1}(\lambda) & \text{if } 0\leq\lambda\leq(1-\alpha)\left(|\hat{h}|^2-1\right)\\ 0 & \text{else,}\end{cases} \tag{58}$$

is a power allocation under power constraint $P(\lambda)=\underset{\hat{h}}{E}\left[p(\hat{h})\right]$. The fact that $\hat{h}\sim\mathcal{CN}(0,1)$ ensures that there exists $\hat{h}_0$, such that $(1-\alpha)\left(|\hat{h}_0|^2-1\right)>0$. Hence, the subsequent achievable rate resulting from the above procedure versus P_{avg} curve can be obtained by varying λ.

Remark 3: While the achievable rate R_- has been proven under weak secrecy constraint, it is worthwhile noting that the same rate is achievable under the variational distance, which is a stronger secrecy metric, as shown in [41]. Note that our converse holds true under the variational distance metric.

IV. Asymptotic Analysis

It is of interest to use Theorem 1 in order to obtain useful insight in some interesting asymptotic cases. Below, we analyze the secrecy capacity at high SNR ($P_{avg}\to\infty$) and at low SNR ($P_{avg}\to 0$), together with the perfect main CSI ($\alpha\to 0$) and no main CSI ($\alpha\to 1$).

A. High-SNR Regime

Our result is summarized in Corollary 1.

Corollary 1: At high SNR, the secrecy capacity C_s^∞ is bounded by:

$$R_-^\infty\leq C_s^\infty\leq R_+^\infty, \tag{59}$$

where R_-^∞ and R_+^∞ are given by:

$$R_-^\infty=\underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\log\left(\left|\frac{h}{g}\right|^2\right)\right] \tag{60}$$

$$R_+^\infty=\underset{\tilde{h},\hat{h}}{E}\left[\log\left(\frac{\left|\sqrt{1-\alpha}\hat{h}+\sqrt{\alpha}\tilde{h}\right|^2}{\left|\tilde{h}\right|^2}\right)\right]^+, \tag{61}$$

and where τ is optimized to maximize R_-.

Proof:

- Asymptotic achievable rate

By Theorem 1, the rate $R_-(\tau)=\underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\log\left(\frac{1+P_0(\tau)|h|^2}{1+P_0(\tau)|g|^2}\right)\right]$ is achievable, for any $\tau\geq 0$. As $P\to\infty$, we have:

$$\lim_{P\to\infty}R_-(\tau)=\lim_{P\to\infty}\underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\log\left(\frac{1+P_0(\tau)|h|^2}{1+P_0(\tau)|g|^2}\right)\right] \tag{62}$$

$$= \underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\lim_{P\to\infty}\log\left(\frac{1+P_0(\tau)|h|^2}{1+P_0(\tau)|g|^2}\right)\right] \quad (63)$$
$$= \underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\log\left(\frac{|h|^2}{|g|^2}\right)\right],$$

where (63) follows from the Dominant Convergence Theorem, since for any P_{avg} value,

$$\left|\log\left(\frac{1+P_0(\tau)|h|^2}{1+P_0(\tau)|g|^2}\right)\right| \leq \left|\log\left(\frac{|h|^2}{|g|^2}\right)\right|,$$

for any $|h|^2 \geq 0$ and $|g|^2 \geq 0$, and

$$\underset{|h|^2,|g|^2,|\hat{h}|^2\geq\tau}{E}\left[\left|\log\left(\frac{|h|^2}{|g|^2}\right)\right|\right] < \infty,$$

since f_g is continuous an bounded, $\left|\int_0^1 \log(x)\,dx\right| = 1$ and $\underset{|h|^2,|\hat{h}|^2\geq\tau}{E}\left[|h|^2\right] \leq \underset{|h|^2,|\hat{h}|^2}{E}\left[|h|^2\right] < \infty$; then the limit in (62) exists and we can insert the limit inside the expectations in (62). To complete this part of the proof, $R_-(\tau)$ is maximized over all $\tau \geq 0$ and the optimum value is achieved at τ_0 that satisfies the necessary optimality condition given by:

$$f_{|\hat{h}|^2}(\tau_0)\left(\underset{|h|^2\,||\hat{h}|^2}{E}\left[\log\left(|h|^2\right) \mid |\hat{h}|^2=\tau_0\right] - \underset{|g|^2}{E}\left[\log\left(|g|^2\right)\right]\right) = 0. \quad (64)$$

- Asymptotic upper bound

For convenience, let UB^{cst} the RHS of (5) where $P\left(\hat{h}\right) = P_{avg}$, a constant power policy independently of $\hat{h}$. The later particular choice provides a lower bound on R_+ and thus:

$$\lim_{P_{avg}\to\infty} R_+ \geq \lim_{P_{avg}\to\infty} UB^{cst}$$
$$= \lim_{P_{avg}\to\infty} \underset{\hat{h},\tilde{h}}{E}\left[\log\left(\frac{1+P_{avg}|\sqrt{1-\alpha}\,\hat{h}+\sqrt{\alpha}\,\tilde{h}|^2}{1+P_{avg}|\tilde{h}|^2}\right)\right]^+$$
$$= \underset{\hat{h},\tilde{h}}{E}\left[\log\left(\frac{|\sqrt{1-\alpha}\,\hat{h}+\sqrt{\alpha}\,\tilde{h}|^2}{|\tilde{h}|^2}\right)\right]^+, \quad (65)$$

where (65) holds by a similar reasoning than the one used to obtain (63). On the other hand, for any $P\left(\hat{h}\right) \geq 0$, the following upper bound on R_+ holds true:

$$R_+ \leq \max_{P(\hat{h})} \underset{\hat{h},\tilde{h}}{E}\left[\log\left(\frac{|\sqrt{1-\alpha}\,\hat{h}+\sqrt{\alpha}\,\tilde{h}|^2}{|\tilde{h}|^2}\right)\right]^+$$
$$= \underset{\hat{h},\tilde{h}}{E}\left[\log\left(\frac{|\sqrt{1-\alpha}\,\hat{h}+\sqrt{\alpha}\,\tilde{h}|^2}{|\tilde{h}|^2}\right)\right]^+. \quad (66)$$

Applying the limit on both sides of (66) completes the proof of our asymptotic upper bound. ■

Clearly, Corollary 1 states that the secrecy capacity is bounded at high SNR confirming that the secret multiplexing gain is equal to zero, regardless of the main channel estimation quality.

B. Low-SNR Regime

While the high-SNR analysis provides somehow a negative result in the sense that the capacity is bounded no matter how P_{avg} increases, we show that at low SNR, and for fading channels with infinite support, the secrecy capacity is asymptotically equal to the capacity of the main channel as if there is no secrecy constraint. Hence, the low-SNR analysis reveals the potential capacity gain provided by partial CSI at the transmitter for any non-null channel estimation quality, i.e., any $\alpha \in [0,1)$. Our result can be stated as follows.

Theorem 2: Assume that the fading h, $\hat{h}$ and g in (1) and (1) have all an infinite support. Let $\alpha \in [0,1)$. Let $C_m\left(P_{avg},\alpha\right)$ be the capacity of the main channel given by $y(i) = h(i)\,x(i) + v(i)$, with CSI-T $\hat{h}_i$ as described by (2) and with CSI-R $\left(h_i,\hat{h}_i\right)$ at time instant i, and with average power constraint P_{avg}. Then the secrecy capacity of the wiretap channel given by (1) satisfies:[1]

$$C_s\left(P_{avg},\alpha\right) \overset{P_{avg}\to 0}{\approx} C_m\left(P_{avg},\alpha\right) \quad (67)$$

Proof: Since the capacity without secrecy constraint cannot be smaller than the one under secrecy constraint, the converse part of Theorem 2 is immediate. To prove the achievability part, let us first consider the main channel as described in Theorem 2. For this channel, let us define the maximum channel gain G by [48]: $G = \sup_{p(x)} \frac{E\left[\gamma_h|x|^2\right]}{E\left[|x|^2\right]}$. It can be verified that because the transmitter has CSI $\hat{h}$, then $G = \sup_{|\hat{h}|^2} E\left[|h|^2 \mid \hat{h}\right] = \sup_{|\hat{h}|^2}\left[(1-\alpha)|\hat{h}|^2+\alpha\right]$. Since by assumption of Theorem 2, $\hat{h}$ has an infinite support, then $G = \infty$. Now, Let us consider the conditional input distribution defined by

$$f_{x\,|\,|\hat{h}|^2}(x \mid |\hat{h}|^2) = \begin{cases}\delta\left(x-\sqrt{P_0}\right) & \text{if} \quad |\hat{h}|^2 \geq \nu,\\ \delta(x) & \text{otherwise,}\end{cases} \quad (68)$$

where $\delta(\cdot)$ is the Dirac delta function, where $P_0 = \frac{P_{avg}}{1-F_{|\hat{h}|^2}(\nu)}$ and where $\nu = \nu\left(P_{avg}\right)$ is a threshold that needs to be determined. Clearly, the input distribution (68) satisfies the power constraint since:

$$E\left[|x|^2\right] = \int_{-\infty}^{+\infty} |x|^2 f_x(x)\,\mathrm{d}x \quad (69)$$
$$= \int_{-\infty}^{+\infty} |x|^2\left(F_{|\hat{h}|^2}(\nu)\,\delta(x) + \left(1-F_{|\hat{h}|^2}(\nu)\right)\delta\left(x-\sqrt{P_0}\right)\right)\mathrm{d}x \quad (70)$$
$$= \left(1-F_{|\hat{h}|^2}(\nu)\right)P_0 \quad (71)$$
$$= P_{avg}. \quad (72)$$

Furthermore, we verify that:

$$\lim_{P_{avg}\to 0} \frac{|E[x]|^2}{E\left[|x|^2\right]} = \lim_{P_{avg}\to 0}\left(1-F_{|\hat{h}|^2}(\nu)\right) \quad (73)$$
$$\lim_{P_{avg}\to 0} \frac{E\left[|h|^2|x|^2\right]}{E\left[|x|^2\right]} = \lim_{P_{avg}\to 0} \underset{|h|^2\;|\hat{h}|^2_{[\nu,\infty)}}{E}[h] \quad (74)$$
$$= \lim_{P_{avg}\to 0}\left[\alpha + (1-\alpha)\frac{\int_\nu^\infty t\,f_{|\hat{h}|^2}(t)\,\mathrm{d}t}{1-F_{|\hat{h}|^2}(\nu)}\right], \quad (75)$$

where $|\hat{h}|^2_{[\nu,\infty)}$ in (74) is the conditional random variable defined by $|\hat{h}|^2_{[\nu,\infty)} = |\hat{h}|^2 \mid |\hat{h}|^2 \geq \nu$. Now, choosing ν such that the limit in (73) is equal to zero and the limit in (74) is equal to G

[1]Please see our formal definition of the notation $\approx$ at the end of Section I.

ensures that the input distribution in (68) is first-order optimal in the sense of [48, Theorem 4]. That is,

$$C_m\left(P_{avg},\alpha\right) \approx \underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right]. \tag{76}$$

The secrecy rate achieved by the above input distribution is given by:

$$R_- = \underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right] - \underset{|g|^2,x}{E}\left[\log\left(1+|g|^2|x|^2\right)\right]. \tag{77}$$

As $P_{avg}\to 0$, the first term in (77) is much larger than the second one as shown below:

$$\lim_{P_{avg}\to 0}\frac{\underset{|g|^2,x}{E}\left[\log\left(1+|g|^2|x|^2\right)\right]}{\underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right]} = \lim_{P_{avg}\to 0}\frac{\frac{\underset{|g|^2,x}{E}\left[\log\left(1+|g|^2|x|^2\right)\right]}{P_{avg}}}{\frac{\underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right]}{P_{avg}}} \tag{78}$$

$$\le \lim_{P_{avg}\to 0}\frac{\frac{\log\left(1+\underset{|g|^2}{E}\left[|g|^2\right]P_{avg}\right)}{P_{avg}}}{\frac{\underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right]}{P_{avg}}} \tag{79}$$

$$= \frac{\underset{|g|^2}{E}\left[|g|^2\right]}{G} \tag{80}$$

$$= 0, \tag{81}$$

where (79) is due to the Jensen's inequality and the independence of x and g, and (80) follows because the input x is first-order optimal. Hence, R_- is asymptotically equal to

$$R_- \approx \underset{|h|^2,x}{E}\left[\log\left(1+|h|^2|x|^2\right)\right]. \tag{82}$$

The rate on the RHS of (82) is asymptotically equal to the capacity of the main channel and hence is the best rate one can achieve. ∎

We emphasize that although in the low-SNR regime, the secrecy capacity is asymptotically equal to the capacity of the main channel as if there is no secrecy constraint, one still needs a wiretap code to guarantee secrecy.

Remark 4: Since the upper bound R_+ cannot be higher than the capacity of the main channel C_m, then a direct application of Theorem 2 establishes that our bounds match at low SNR. That is, $R_-\left(P_{avg},\alpha\right) \overset{P_{avg}\to 0}{\approx} R_+\left(P_{avg},\alpha\right)$, for any $\alpha\in[0,1)$, including the perfect main CSI-T ($\alpha=0$) treated in [25].

C. Perfect And No Main CSI Extremes

When specialized to the no main CSI case, the lower and the upper bounds in Theorem 1 coincide, providing a trivial secrecy capacity. On the other hand, in case of perfect main CSI, the upper bound coincides with the secrecy capacity of a wiretap fading channel under the assumption of asymptotically long coherence intervals derived in [17, Theorem 2]. Our result is formalized in Corollary 2.

Corollary 2: In Case of no main CSI, the secrecy capacity C_s is equal to zero: $C_s=0$; whereas, in case of perfect main CSI, the lower and the upper bounds in (5) reduce to:

$$R_-^{\text{perfect}} \le C_s \le R_+^{\text{perfect}}, \tag{85}$$

where R_-^{perfect} and R_+^{perfect} are given by:

$$R_-^{\text{perfect}} = \underset{|h|^2\ge\tau_0,|g|^2}{E}\left[\log\left(\frac{1+P_{avg}e^{\tau}|h|^2}{1+P_{avg}e^{\tau}|g|^2}\right)\right] \tag{86}$$

$$R_+^{\text{perfect}} = \max_{P(h)}\underset{|h|^2,|g|^2}{E}\left[\log\left(\frac{1+P(h)|h|^2}{1+P(h)|g|^2}\right)\right]^+, \tag{87}$$

and where τ_0 is a solution of:

$$\underset{|h|^2\ge\tau}{E}\left[\frac{P_{avg}e^{\tau}|h|^2}{1+P_{avg}e^{\tau}|h|^2}\right] - \underset{|g|^2}{E}\left[\frac{P_{avg}e^{\tau}|g|^2}{1+P_{avg}e^{\tau}|g|^2}\right]\left(1-F_{|\hat{h}|^2}(\tau)\right)$$
$$-f_{|h|^2}(\tau)\left(\log(1+P\tau e^{\tau}) - \underset{|g|^2}{E}\left[\log\left(1+Pe^{\tau}|g|^2\right)\right]\right) = 0. \tag{88}$$

Proof:

- No Main CSI

In this case, we have $\alpha=1$, i.e., $h=\tilde{h}$ and $P\left(\hat{h}\right)=P$. The upper bound in (5) is equal to zero, and so is the secrecy capacity.

- Perfect Main CSI

In this case, we have $\alpha=0$, i.e., $h=\hat{h}$ and $P\left(\hat{h}\right)=P(h)$. Since g is independent of x, then $\left(h',g'\right)=(h,g)$ is a valid choice that is when applied to (50) provides the upper bound in Corollary 2. The lower bound follows by specializing the result in Theorem 1 to the case where $h=\hat{h}$. ∎

Remark 5: In case of no main CSI, the secrecy capacity is not zero solely because of the unavailability of the main CSI-T, but also because the main and the eavesdropper channels have identical statistics. Should the main channel be better than the eavesdropper channel on average, one may still be able to achieve a positive secrecy rate, even without main CSI-T.

V. Application: i.i.d. Rayleigh Fading Channels

In this section, we apply the results derived in the previous sections to i.i.d. Rayleigh fading channels. That is, the main channel h and the eavesdropper channel g are circularly symmetric complex Gaussian with mean zero and variance one, and so are $\hat{h}$ and $\tilde{h}$. In our derivations summarized in Table I, we have used the fact that $|h|^2$, $|\hat{h}|^2$ and $|g|^2$ are all exponentially distributed with common pdf: $f_X(x)=e^{-x}$, where the subscript X is any of the r.v. $|h|^2$, $|\hat{h}|^2$ and $|g|^2$. Additionally, the conditional r.v. $|h|^2\mid\hat{h}$ has a PDF given by:

$$f_{|h|^2|\hat{h}}(|h|^2\mid\hat{h}) = \frac{1}{\alpha}e^{-\frac{|h|^2+(1-\alpha)|\hat{h}|^2}{\alpha}}I_0\left(2\sqrt{\frac{(1-\alpha)|\hat{h}|^2|h|^2}{\alpha^2}}\right), \tag{89}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind. Also, if $\hat{\theta}$ and $\tilde{\theta}$ are uniformly distributed between $[-\pi,\pi]$, then the PDF of $\theta=\left(\hat{\theta}-\tilde{\theta}\right)$ is:

$$f_\theta(\theta) = \begin{cases}\frac{1}{(2\pi)^2}(2\pi+\theta) & -2\pi\le\theta<0\\ \frac{1}{(2\pi)^2}(2\pi-\theta) & 0\le\theta<2\pi\\ 0 & \text{elsewhere.}\end{cases} \tag{90}$$

Similarly, if $\hat{\rho}$ and $\tilde{\rho}$ are Rayleigh distributed, with distribution $f_{\hat{\rho}}(\hat{\rho}) = 2\hat{\rho}\,e^{-\hat{\rho}^2} = f_{\tilde{\rho}}(\hat{\rho})$, then their ratio $\rho=\hat{\rho}/\tilde{\rho}$ has the PDF:

$$f_\rho(\rho) = \frac{2\rho}{(1+\rho^2)^2}, \tag{91}$$

for $\rho \geq 0$. Finally, for a Rayleigh fading channel, we have: $P_0(\tau) = P_0'(\tau) = P_{avg} e^{\tau}$. In Tab. I, $\mathrm{E_i}(x) = \int_x^{\infty} \frac{e^{-t}}{t} \mathrm{d}t$ is the exponential integral function and $\gamma \approx 0.577216$ is the Euler's constant, whereas LHS stands for left hand side.

In order to illustrate the result at low SNR, we note that for the main channel as described by Theorem 2, we cannot apply directly the low-SNR characterization in [49] since therein, the authors assume that the receiver knows both $\hat{h}$ and $\tilde{h}$. Nevertheless, by considering $\nu = \log\left(P_{avg}\right) - 2\log\log\left(\frac{1}{P_{avg}}\right)$ in (68), one can verify that the limits in (73) and (74) are equal to 0 and ∞, respectively; and thus the input distribution given by (68), with such a choice of ν is first-order optimal. Furthermore, applied to the main channel, this input distribution achieves a rate R_m equal to:

$$
\begin{aligned}
R_m &= \underset{|h|^2,x}{E}\left[\log\left(1+|h|^2\,|x|^2\right)\right] && (92)\\
&= \mathrm{Prob}\left\{|\hat{h}|^2 \geq \nu\right\} \underset{|h|^2,|\hat{h}|^2}{E}\left[\log\left(1+|h|^2\,P_0\right) \mid |\hat{h}|^2 \geq \nu\right] && (93)\\
&\approx \mathrm{Prob}\left\{|\hat{h}|^2 \geq \nu\right\} P_0 \underset{|h|^2,|\hat{h}|^2}{E}\left[|h|^2 \mid |\hat{h}|^2 \geq \nu\right] && (94)\\
&= P_{avg} \underset{|\hat{h}|^2}{E}\left[(1-\alpha)\,|\hat{h}|^2 + \alpha \mid |\hat{h}|^2 \geq \nu\right] && (95)\\
&= (1-\alpha)\;P_{avg} \underset{|\hat{h}|^2}{E}\left[|\hat{h}|^2 \mid |\hat{h}|^2 \geq \nu\right] + \alpha\;P_{avg} && (96)\\
&\geq (1-\alpha)\;P_{avg}\;\nu + \alpha\;P_{avg} && (97)\\
&\approx (1-\alpha)\;P_{avg}\;\nu && (98)\\
&\approx (1-\alpha)\;P_{avg}\;\log\left(\frac{1}{P_{avg}}\right) && (99)
\end{aligned}
$$

where (94) follows using the fact that $\log(1+x) \approx x$ when $x \to 0$ along with the fact that $P_0 = \frac{1}{\log^2\left(P_{avg}\right)}$ which converges to 0 as $P_{avg} \to 0$; (95) holds true since given $\hat{h}$, the mean of $|h|^2$ is equal to $(1-\alpha)\,|\hat{h}|^2 + \alpha$; and (98) is true because ν converges to ∞ as $P_{avg} \to 0$. The RHS of (99) is the asymptotic capacity of an enhanced main channel where the transmitter still knows $\hat{h}$, but the receiver now knows $\left(h_i, \overline{h}_i, \hat{h}_i\right)$. The capacity of this enhanced channel –an upper bound on the capacity of our main channel– is asymptotically equal to $(1-\alpha)\;P_{avg}\;\log\left(\frac{1}{P_{avg}}\right)$ as it has been shown in [49]. We conclude that the asymptotic capacity of the main channel described in Theorem 2 at low SNR for Rayleigh fading channels is given by:

$$
C_m\left(P_{avg}, \alpha\right) \approx (1-\alpha)\;P_{avg}\;\log\left(\frac{1}{P_{avg}}\right). \qquad (100)
$$

Applying Theorem 2, we immediately have:

$$
C_s\left(P_{avg}, \alpha\right) \approx (1-\alpha)\;P_{avg}\;\log\left(\frac{1}{P_{avg}}\right). \qquad (101)
$$

Note that to establish (100), we have utilized an enhanced main channel so that known results at low SNR [49] can be exploited in our proof. Our result however, is applicable to the main channel as described in Theorem 2 where the receiver knows h_i and $\overline{h}_i$, but not $\hat{h}_i$. This suggests that the (somehow strong) assumption by which the receiver knows both $\hat{h}$ and $\tilde{h}$ as required in [49] is apparently not fundamental in the regime of interest.

VI. Numerical Results

In this section, numerical results are provided for i.i.d. Rayleigh fading channels. Figure 1 depicts the lower and the upper bounds in Theorem 1 in nats per channel use (npcu) versus P_{avg} designated as SNR, for different main channel estimation error variances α. Also shown in Fig. 1 are the high-SNR bounds given by (59) along with the corresponding bounds to perfect and no-main CSI extremes given by (85). As can be seen in Fig. 1, the secrecy rate is strictly positive even for a poor main channel estimation quality ($\alpha = 0.9$). Although there is a gap between the lower and the upper bounds for all $\alpha \in [0, 1)$, this gap is bounded for all SNR values.

Figure 2 assesses the rate loss incurred by the proposed on-off power scheme, compared to the one obtained from KKT condition in Remark 2, for $\alpha = 0.5$. As can be seen in Fig. 2, the procedure described in Remark 2, albeit complex and time-consuming, does not provide a substantial gain, in the setting considered in our paper. Indeed, the rate achieved by the proposed on-off power policy and the one resulting from the KKT condition are very close for all the SNR range displayed. This observation holds true for different values of α.

Figure 3 depicts the bounds versus α for different SNR values. The bounds match for $\alpha = 1$, confirming that the secrecy capacity in case of no main CSI is equal to zero.

In Fig. 4, the optimal values of τ_0 versus SNR is displayed. The curves in Fig. 4 have been obtained by solving the necessary condition (43) or its high-SNR version (88). Interestingly, at high SNR, and for a given channel estimation error α, τ_0 converges to a fixed value, say $\tau_0^{\infty}(\alpha)$, which suggests that at high SNR, if the transmitter is provided this value, it would be able to achieve the same secrecy rate without the need of $\hat{h}$. Note also that for a given SNR value, τ_0 decreases with the channel estimation quality.

In Fig. 5 and Fig. 6, we present results in the low-SNR regime for an estimation error equal to $\alpha = 0.5$ and for the perfect CSI-T ($\alpha = 0$) case, respectively. In addition to the upper and the lower bounds, we have also depicted the capacity of the main channel with noisy CSI-T along with the asymptotic expression given by (100). The curve corresponding to the capacity of the main channel has been obtained by evaluation of [46]:

$$
C_m = \max_{\gamma} E\left[\log\left(1+\gamma\left(\hat{h}\right)\;|h|^2\right)\right], \qquad (102)
$$

where the expectation is with respect to the joint distribution of $|h|^2$ and $\hat{h}$ and where the maximization is over all power policies that satisfy the power constraint. In Fig. 5, We note that as SNR decreases, our bounds asymptotically match in agreement with our discussion in Section IV. Both bounds get closer to the capacity of the main channel (without secrecy constraint) although the convergence seems slow, and one expects the three curves to match at SNR values below -40 dB. The low-SNR characterization in [49] seems slightly optimistic and is more accurate for strictly positive values of α.

VII. Conclusion

The secrecy capacity of i.i.d. fast fading channels, under imperfect main channel estimation at the transmitter,

Table I
SUMMARY OF THE DERIVED RESULTS FOR I.I.D. RAYLEIGH FADING CHANNELS.

General i.i.d.	i.i.d. Rayleigh
(6)	$\int_{\hat{h}=\tau_0}^{\infty}\int_{h=0}^{\infty}\int_{g=0}^{\infty}\log\left(\frac{1+P_0(\tau_0)h}{1+P_0(\tau_0)g}\right)e^{-g}f_{\lvert h\rvert^2\lvert\hat{h}}\left(h\mid\sqrt{\hat{h}}\right)e^{-\hat{h}}\,\mathrm{d}g\,\mathrm{d}h\,\mathrm{d}\hat{h}$
(7)	$\int_{\hat{\rho}=0}^{\infty}\int_{\tilde{\theta}=-\pi}^{\pi}\int_{\tilde{\rho}=0}^{\frac{\hat{\rho}}{\rho_0(\tilde{\theta})}}\log\left(\frac{1+P(\hat{\rho})\left((1-\alpha)\hat{\rho}^2+\alpha\tilde{\rho}^2+2\sqrt{\alpha(1-\alpha)}\hat{\rho}\tilde{\rho}\cos(\tilde{\theta})\right)}{1+P(\hat{\rho})\tilde{\rho}^2}\right)f_{\tilde{\rho}}(\tilde{\rho})\,\frac{1}{2\pi}\,f_{\hat{\rho}}(\hat{\rho})\,\mathrm{d}\tilde{\rho}\,\mathrm{d}\tilde{\theta}\,\mathrm{d}\hat{\rho}$
(15)	$\int_{\hat{h}=\tau_0}^{\infty}\int_{h=0}^{\infty}\int_{g=0}^{\infty}\frac{P_0(\tau_0)h}{(1+P_0(\tau_0)h)}f_{\lvert h\rvert^2\lvert\hat{h}}\left(h\mid\sqrt{\hat{h}}\right)e^{-\hat{h}}\,\mathrm{d}g\,\mathrm{d}h\,\mathrm{d}\hat{h}$ $= e^{-\tau_0}\left(1-\frac{e^{\frac{1}{P_0(\tau_0)}}\,\mathrm{E_i}\left(\frac{1}{P_0(\tau_0)}\right)}{P_0(\tau_0)}-e^{\frac{1}{P_0(\tau_0)}}\,\mathrm{E_i}\left(\frac{1}{P_0(\tau_0)}\right)+\int_{h=0}^{\infty}\log\left(1+P_0(\tau_0)h\right)f_{\lvert h\rvert^2\lvert\hat{h}}\left(h\mid\sqrt{\hat{h}}\right)\mathrm{d}h\right)$
(52)	$\int_{\tilde{\theta}=-\pi}^{\pi}\int_{\tilde{\rho}=0}^{\frac{\rho}{\rho_0(\tilde{\theta})}}\left(\frac{(1-\alpha)\hat{\rho}^2+\alpha\tilde{\rho}^2+2\sqrt{\alpha(1-\alpha)}\hat{\rho}\tilde{\rho}\cos(\tilde{\theta})}{1+P(\hat{\rho})\left((1-\alpha)\hat{\rho}^2+\alpha\tilde{\rho}^2+2\sqrt{\alpha(1-\alpha)}\hat{\rho}\tilde{\rho}\cos(\tilde{\theta})\right)}-\frac{\tilde{\rho}^2}{1+P(\hat{\rho})\tilde{\rho}^2}\right)f_{\tilde{\rho}}(\tilde{\rho})\,\frac{1}{2\pi}\mathrm{d}\tilde{\rho}\,\mathrm{d}\tilde{\theta}-\lambda=0$
(60)	$e^{-\tau_0}\left(\gamma+e^{\tau_0}\,\mathrm{E_i}(\tau_0)-e^{\tau_0}\,\mathrm{E_i}\left(\frac{\tau_0}{\alpha}\right)+\mathrm{E_i}\left(\frac{1-\alpha}{\alpha}\tau_0\right)+\log\left((1-\alpha)\tau_0\right)\right)$
(61)	$\int_{-2\pi}^{2\pi}\int_{\frac{\hat{\rho}}{\rho_0(\theta)}}^{\infty}\log\left((1-\alpha)\rho^2+\alpha+2\sqrt{\alpha(1-\alpha)}\rho\cos(\theta)\right)f_{\boldsymbol{\rho}}(\rho)\,f_{\boldsymbol{\theta}}(\theta)$
(64)	$\mathrm{E_i}\left(\frac{1-\alpha}{\alpha}\tau_0\right)+\log(\tau_0)+\log(1-\alpha)+\gamma=0$
(86)	$e^{-\tau_0}\left(e^{\frac{e^{-\tau_0}}{P_{avg}}}\left(e^{\tau_0}\,\mathrm{E_i}\left(\frac{e^{-\tau_0}}{P_{avg}}+\tau_0\right)-\mathrm{E_i}\left(\frac{e^{-\tau_0}}{P_{avg}}\right)\right)+\log\left(1+\tau_0e^{\tau_0}P_{avg}\right)\right)$
(87)	$\max_{P(h)}\int_0^{\infty}\left(\log(1+hP(h))-e^{\frac{1}{P(h)}}\left(\mathrm{E_i}\left(\frac{1}{P(h)}\right)-\mathrm{E_i}\left(h+\frac{1}{P(h)}\right)\right)\right)e^{-h}\,\mathrm{d}h$
(88)	$\frac{e^{-2\tau}}{P_{avg}}\left(e^{\frac{e^{-\tau}}{P_{avg}}}\left(\left(1+e^{\tau}P_{avg}\right)\mathrm{E_i}\left(\frac{e^{-\tau}}{P_{avg}}\right)-e^{\tau}\,\mathrm{E_i}\left(\frac{e^{-\tau}}{P_{avg}}+\tau\right)\right)-e^{\tau}P_{avg}\log\left(1+e^{\tau}P_{avg}\tau\right)\right)=0$

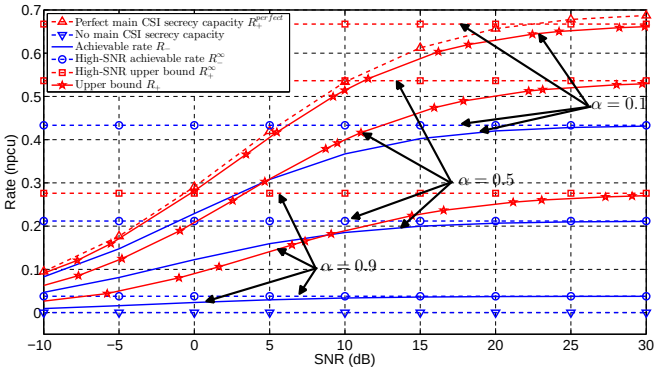

Figure 1. Achievable rate and upper bound for i.i.d. Rayleigh fading channels, with various main channel estimation errors α.

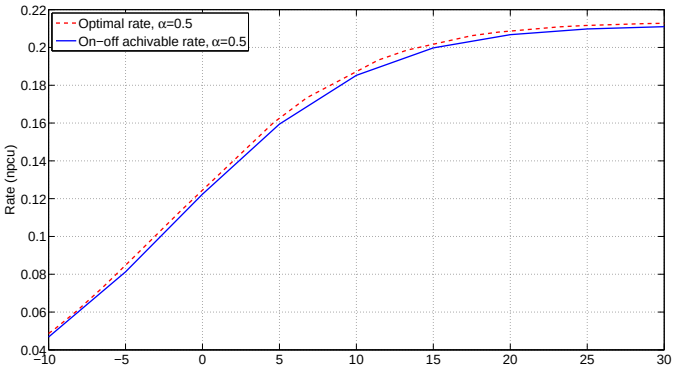

Figure 2. On-off achievable rate and the one resulting from the KKT condition versus SNR, for $\alpha = 0.5$.

is addressed. Lower and upper bounds are derived for a given channel estimation quality, and the gap between these bounds is characterized numerically. In addition, special cases regarding perfect and no main CSI at the transmitter are studied. Particularly, it is shown that that our bounds coincide with recently derived bounds for the i.i.d. fading channels. Furthermore, insightful asymptotic analyses at high SNR and at low SNR are provided. Perhaps surprising, it is found that the secrecy capacity is equal to the capacity of the main channel without secrecy constraint at asymptotically low SNR.

Our framework shows, for instance, that even a poor main channel estimator at the transmitter can help establish secure communication. This fact has also been demonstrated in e.g. [41], although in a slightly different setting. Furthermore, a simple constant rate on-off power scheme is enough to achieve a positive secrecy rate. It is to be reminded that one can enhance the later achievable secrecy rate by optimizing the transmit power with respect to the main channel estimation, at the expense of increasing the system complexity.

Finally, we note that our upper bound relies on the eavesdropper's channel having the same statistics as the main channel estimation error. This leaves open the problem of determining a generic upper bound.

ACKNOWLEDGMENT

The authors would like to thank the editor Dr. Andrew Thangaraj for volunteering his time to handle this paper and the anonymous reviewers for their valuable comments that

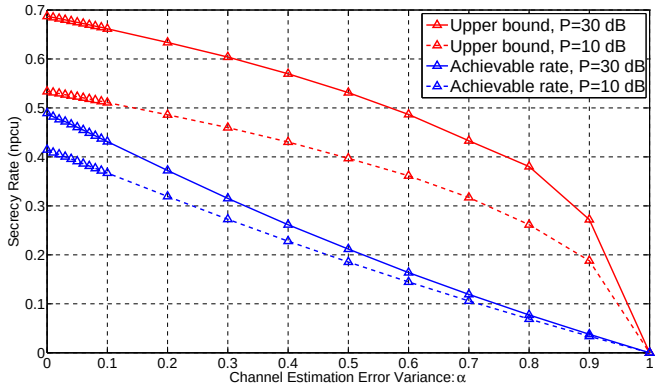

Figure 3. Achievable rate and upper bound for i.i.d. Rayleigh fading channels in function of α.

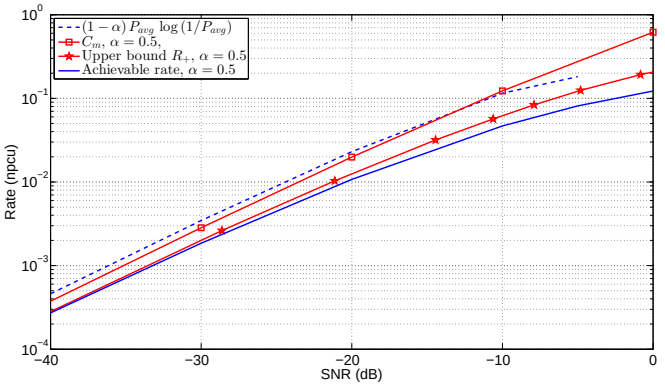

Figure 5. Achievable rate, upper bound, capacity of the main channel and the asymptotic expression given by (100) for i.i.d. Rayleigh fading channels, with estimation error $\alpha = 0.5$.

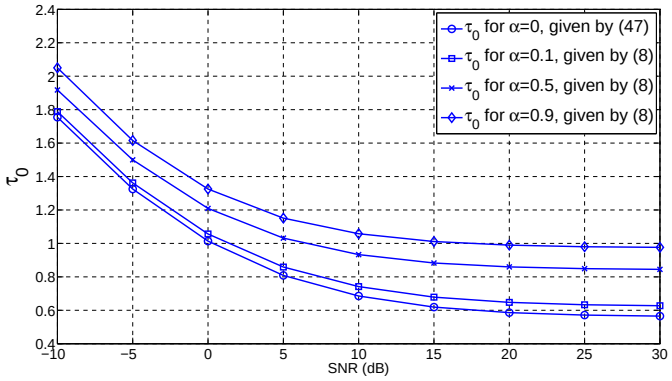

Figure 4. Optimal on-off power parameter τ_0 versus SNR for i.i.d. Rayleigh fading channels and for various values of α.

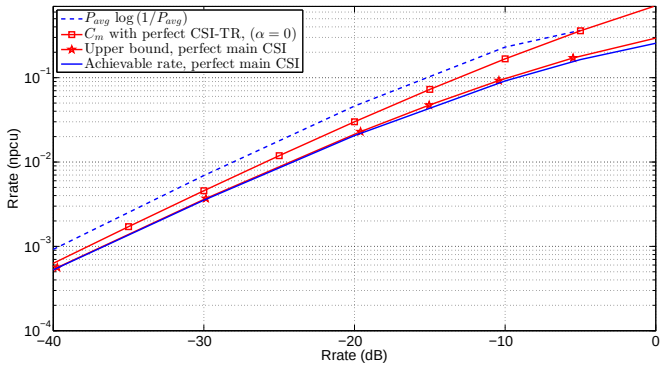

Figure 6. Achievable rate, upper bound, capacity of the main channel and the asymptotic expression given by (100) for i.i.d. Rayleigh fading channels, with perfect CSI-T, $\alpha = 0$.

have enhanced the technical quality and the lucidity of this paper.

References

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

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

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

[4] A. Thangaraj, S. Dihidar, A. Calderbank, S. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.

[5] A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using large-girth ldpc codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 585–594, Sep. 2011.

[6] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, Oct 2011.

[7] O. Koyluoglu and H. El-Gamal, "Polar coding for secure transmission and key agreement," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1472–1483, Oct. 2012.

[8] I. Tal and A. Vardy, "Channel upgrading for semantically-secure encryption on wiretap channels," in *Proc. IEEE International Symposium onInformation Theory (ISIT'2013)*, Istanbul, Turkey, Jul. 2013, pp. 1561–1565.

[9] M. Andersson, R. Schaefer, T. Oechtering, and M. Skoglund, "Polar coding for bidirectional broadcast channels with common and confidential messages," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1901–1908, Sep. 2013.

[10] J. C. Belfiore and F. Oggier, "Secrecy gain: A wiretap lattice code design," in *Proc. IEEE 2010 International Symposium on Information Theory and its Applications (ISITA)*, Taichung, Taiwan, Oct. 2010, pp. 174–178.

[11] X. He and A. Yener, "Providing secrecy with structured codes: Two-user Gaussian channels," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2121–2138, Apr. 2014.

[12] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Advances in Cryptology - CRYPTO 2012*, ser. Lecture Notes in Computer Science, R. Safavi-Naini and R. Canetti, Eds. Springer Berlin Heidelberg, 2012, vol. 7417, pp. 294–311.

[13] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proc. International Symposium on Information Theory (ISIT'2005)*, Adelaide, Australia, Sept. 2005, pp. 2152–2155.

[14] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. International Symposium on Information Theory (ISIT'2006)*, Seattle, Washington, USA, Jul. 2006, pp. 356–360.

[15] 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.

[16] P. Wang, G. Yu, and Z. Zhang, "On the secrecy capacity of fading wireless channel with multiple eavesdroppers," in *Proc. International Symposium on Information Theory (ISIT'2007)*, Nice, France, Sep. 2007, pp. 1301–1305.

[17] 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.

[18] 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.

[19] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2453 –2469, Jun. 2008.

[20] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *Proc. 41st Annual Conference on Information Sciences and Systems (CISS'2007)*, The John Hopkins University, Baltimore, Maryland, USA, March 2007, pp. 905–910.

[21] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 4033–4039, 2009.

[22] A. Khisti and G. Wornell, "Secure transmission with multiple antennas. Part II: The MIMOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[23] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[24] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[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] I. Hero, A.O., "Secure space-time communication," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3235–3249, December 2003.

[27] M. Yuksel and E. Erkip, "Diversity-multiplexing tradeoff for the multiple-antenna wire-tap channel," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1 –10, Jan. 2011.

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

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] ——, "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.

[43] U. M. Maurer and S.Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *in Proc. of the EUROCRYPT 2000, on Advances in Cryptology*, vol. 1807, 2000, pp. 352–368.

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

[45] M. Bloch and J. Laneman, "Strong secrecy from channel resolvability," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[46] 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.

[47] J. Thomas, "Feedback can at most double gaussian multiple access channel capacity (corresp.)," *IEEE Trans. Inform. Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.

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

[49] S. Borade and L. Zheng, "Wideband fading channels with feedback," *IEEE Trans. Inform. Theory*, vol. 56, no. 12, pp. 6058–6065, Dec. 2010.

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.