On the Secrecy Capacity of Fading Channels

Praveen Kumar Gopala, Lifeng Lai and Hesham El Gamal
Department of Electrical and Computer Engineering
The Ohio State University
Columbus, OH 43210
Email: {gopalap,lail,helgamal}@ece.osu.edu

Abstract

We consider the secure transmission of information over an ergodic fading channel in the presence of an eavesdropper. Our eavesdropper can be viewed as the wireless counterpart of Wyner's wiretapper. The secrecy capacity of such a system is characterized under the assumption of asymptotically long coherence intervals. We first consider the full Channel State Information (CSI) case, where the transmitter has access to the channel gains of the legitimate receiver and the eavesdropper. The secrecy capacity under this full CSI assumption serves as an upper bound for the secrecy capacity when only the CSI of the legitimate receiver is known at the transmitter, which is characterized next. In each scenario, the perfect secrecy capacity is obtained along with the optimal power and rate allocation strategies. We then propose a low-complexity on/off power allocation strategy that achieves near-optimal performance with only the main channel CSI. More specifically, this scheme is shown to be asymptotically optimal as the average SNR goes to infinity, and interestingly, is shown to attain the secrecy capacity under the full CSI assumption. Remarkably, our results reveal the positive impact of fading on the secrecy capacity and establish the critical role of rate adaptation, based on the main channel CSI, in facilitating secure communications over slow fading channels.

1 Introduction

The notion of information-theoretic secrecy was first introduced by Shannon [1]. This strong notion of secrecy does not rely on any assumptions on the computational resources of the eavesdropper. More specifically, perfect information-theoretic secrecy requires that $I(W;Z) = 0$, i.e., the signal Z received by the eavesdropper does not provide any additional information about the transmitted message W. Shannon considered a scenario where both the legitimate receiver and the eavesdropper have direct access to the transmitted signal. Under this model, he proved a negative result implying that the achievability of perfect secrecy requires the entropy of the private key K, used to encrypt the message W, to be larger than or equal to the entropy of the message itself (i.e., $H(K) \geq H(W)$ for perfect secrecy). However, it was later shown by Wyner in [2] that this negative result was a consequence of the over-restrictive model used in [1]. Wyner introduced the wiretap channel which accounts for the difference in the two noise processes, as observed by the destination and wiretapper. In this model, the wiretapper has no computational limitations and is assumed to know the codebook used by the transmitter. Under the assumption that the wiretapper's signal is a degraded version of the destination's signal, Wyner characterized the tradeoff between the information rate to the destination and the level of ignorance at the wiretapper (measured by its equivocation), and showed that it is possible to achieve a non-zero secrecy capacity. This work was later extended to non-degraded channels by Csiszár and Körner [3], where it was shown that if the main channel is less noisy or more capable than the wiretapper channel, then it is possible to achieve a non-zero secrecy capacity.

More recently, the effect of slow fading on the secrecy capacity was studied in [8, 9]. In these works, it is assumed that the fading is quasi-static which leads to an alternative definition of outage probability, wherein secure communications can be guaranteed only for the fraction of time when the main channel is stronger than the channel seen by the eavesdropper. This performance metric appears to have an operational significance only in delay sensitive applications with full Channel State Information (CSI). The absence of CSI sheds doubt on the operational significance of outage-based secrecy since it limits the ability of the source to know which parts of the message are decoded by the eavesdropper. In this paper, we focus on delay-tolerant applications which allow for the adoption of an ergodic version of the slow fading channel, instead of the outage-based formulation. Quite interestingly, we show in the sequel that, under this model, one can achieve a perfectly secure non-zero rate even when the eavesdropper channel is more capable then the legitimate channel **on the average**. In particular, our work here characterizes the secrecy capacity of the slow fading channel in the presence of an eavesdropper [10]. Our eavesdropper is the wireless counterpart of Wyner's wiretapper. We first assume that the transmitter knows the CSI of both the legitimate and eavesdropper channels, and derive the optimal power allocation strategy that achieves the secrecy capacity. Next we consider the case where the transmitter only knows the legitimate channel CSI and, again, derive the optimal power allocation strategy. We then propose an on/off power transmission scheme, with variable rate allocation, which approaches the optimal performance for asymptotically large average SNR. Interestingly, this scheme is also shown to attain the secrecy capacity under the full CSI assumption which implies that, at high SNR values, the additional knowledge of the eavesdropper CSI does not yield any gains in terms of the secrecy capacity for slow fading channels. Finally, our theoretical and numerical results are used to argue that rate adaptation plays a more critical role than power control in achieving the secrecy capacity of slow fading channels. This observation contrasts the scenario without secrecy constraints, where transmission strategies with constant rate are able to achieve capacity [4].

2 System Model

The system model is illustrated in Fig. 1. The source S communicates with a destination D in the presence of an eavesdropper E. During any coherence interval i, the signal received by the destination and the eavesdropper are given by, respectively

$$\begin{aligned} y(i) &= g_M(i)x(i) + w_M(i), \\ z(i) &= g_E(i)x(i) + w_E(i), \end{aligned}$$

where $g_M(i), g_E(i)$ are the channel gains from the source to the legitimate receiver (main channel) and the eavesdropper (eavesdropper channel) respectively, and $w_M(i), w_E(i)$ represent the i.i.d additive Gaussian noise with unit variance at the destination and the eavesdropper respectively.

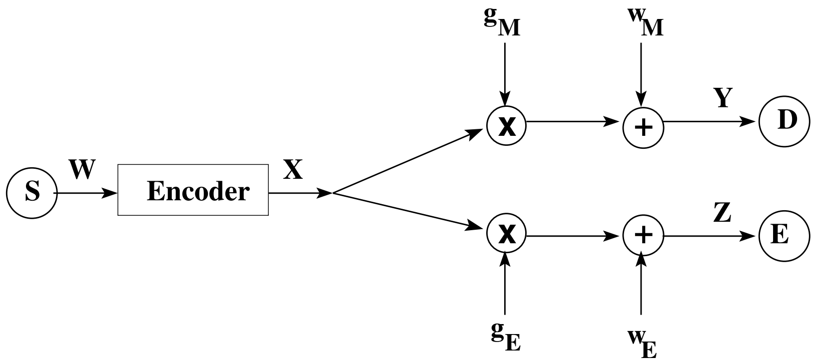

Figure 1: The Fading Channel with an Eavesdropper

We denote the fading power gains of the main and eavesdropper channels by $h_M(i) = |g_M(i)|^2$ and $h_E(i) = |g_E(i)|^2$ respectively. We assume that both channels experience block fading, where the channel gains remain constant during each coherence interval and change independently from one coherence interval to the next. The fading process is assumed to be ergodic with a bounded continuous distribution. Moreover, the fading coefficients of the destination and the eavesdropper in any coherence interval are assumed to be independent of each other. We further assume that the number of channel uses n_1 within each coherence interval is large enough to allow for invoking random coding arguments. As shown in the sequel, this assumption is instrumental in our achievability proofs.

The source wishes to send a message $W \in \mathcal{W} = \{1, 2, \cdots, M\}$ to the destination. An (M, n) code consists of the following elements: 1) a stochastic encoder $f_n(.)$ at the source that maps the message[1] w to a codeword $x^n \in \mathcal{X}^n$, and 2) a decoding function ϕ: $\mathcal{Y}^n \rightarrow \mathcal{W}$ at the legitimate receiver. The average error probability of an (M, n) code at the legitimate receiver is defined as

$$P_e^n = \sum_{w \in \mathcal{W}} \frac{1}{M} \Pr(\phi(y^n) \neq w | w \text{ was sent}). \tag{1}$$

The equivocation rate R_e at the eavesdropper is defined as the entropy rate of the transmitted message conditioned on the available CSI and the channel outputs at the eavesdropper, i.e.,

$$R_e \triangleq \frac{1}{n} H(W|Z^n, h_M^n, h_E^n) , \tag{2}$$

where $h_M^n = \{h_M(1), \cdots, h_M(n)\}$ and $h_E^n = \{h_E(1), \cdots, h_E(n)\}$ denote the channel power gains of the legitimate receiver and the eavesdropper in n coherence intervals, respectively. It indicates the level of ignorance of the transmitted message W at the eavesdropper. In this paper we consider only perfect secrecy which requires the equivocation rate R_e to be equal to the message rate. The perfect secrecy rate R_s is said to be achievable if for any $\epsilon > 0$, there exists a sequence of codes $(2^{nR_s}, n)$ such that for any $n \geq n(\epsilon)$, we have

$$\begin{aligned} P_e^n &\leq \epsilon, \\ R_e &= \frac{1}{n} H(W|Z^n, h_M^n, h_E^n) \geq R_s - \epsilon. \end{aligned}$$

[1]The realizations of the random variables W, X, Y, Z are represented by w, x, y, z respectively in the sequel.

The secrecy capacity C_s is defined as the maximum achievable perfect secrecy rate, i.e.,

$$C_s \triangleq \sup_{P_e^n \leq \epsilon} R_s \ . \tag{3}$$

Throughout the sequel, we assume that the CSI is known at the destination perfectly. Based on the available CSI, the transmitter adapts its transmission power **and rate** to maximize the perfect secrecy rate subject to a long-term average power constraint $\bar{P}$.

3 Full CSI at the Transmitter

Here we assume that at the beginning of each coherence interval, the transmitter knows the channel states of the legitimate receiver and the eavesdropper perfectly. When h_M and h_E are both known at the transmitter, one would expect the optimal scheme to allow for transmission only when $h_M > h_E$, and to adapt the transmitted power according to the instantaneous values of h_M and h_E. The following result formalizes this intuitive argument.

Theorem 1 *When the channel gains of both the legitimate receiver and the eavesdropper are known at the transmitter, the secrecy capacity is given by*

$$C_s^{(F)} = \max_{P(h_M,h_E)} \int_0^\infty \int_{h_E}^\infty \Big[\log\left(1 + h_M P(h_M,h_E)\right) - \log\left(1 + h_E P(h_M,h_E)\right) \Big] f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E, \tag{4}$$

$$\textit{such that} \qquad \mathbb{E}\{P(h_M,h_E)\} \ \leq \ \bar{P}. \tag{5}$$

Proof: A detailed proof of achievability and the converse part is provided in the Appendix. Here, we outline the scheme used in the achievability part. In this scheme, transmission occurs only when $h_M > h_E$, and uses the power allocation policy $P(h_M,h_E)$ that satisfies the average power constraint (5). Moreover, the codeword rate at each instant is set to be $\log\left(1 + h_M P(h_M,h_E)\right)$, which varies according to the instantaneous channel gains. The achievable perfect secrecy rate at any instant is then given by $[\log\left(1 + h_M P(h_M,h_E)\right) - \log\left(1 + h_E P(h_M,h_E)\right)]^+$. Averaging over all fading realizations, we get the average achievable perfect secrecy rate as

$$\begin{aligned} R_s^{(F)} &= \iint [\log\left(1 + h_M P(h_M,h_E)\right) - \log\left(1 + h_E P(h_M,h_E)\right)]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \\ &= \int_0^\infty \int_{h_E}^\infty \Big[\log\left(1 + h_M P(h_M,h_E)\right) - \log\left(1 + h_E P(h_M,h_E)\right) \Big] f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \ . \end{aligned}$$

One can then optimize over all feasible power control policies $P(h_M,h_E)$ to maximize the perfect secrecy rate. □

We now derive the optimal power allocation policy that achieves the secrecy capacity under the full CSI assumption. It is easy to check that the objective function is concave in $P(h_M,h_E)$, and hence, by using the Lagrangian maximization approach for solving (4), we get the following optimality condition

$$\frac{\partial R_s^{(F)}}{\partial P(h_M,h_E)} \;=\; \frac{h_M}{1+h_M P(h_M,h_E)} - \frac{h_E}{1+h_E P(h_M,h_E)} - \lambda \;=\; 0,$$

whose solution is

$$P(h_M,h_E) \;=\; \frac{1}{2}\left[\sqrt{\left(\frac{1}{h_E}-\frac{1}{h_M}\right)^2+\frac{4}{\lambda}\left(\frac{1}{h_E}-\frac{1}{h_M}\right)}-\left(\frac{1}{h_M}+\frac{1}{h_E}\right)\right]. \tag{6}$$

If for some (h_M, h_E), the value of $P(h_M, h_E)$ obtained from (6) is negative, then it follows from the concavity of the objective function w.r.t. $P(h_M, h_E)$ that the optimal value of $P(h_M, h_E)$ is 0. Thus the optimal power allocation policy at the transmitter is given by

$$P(h_M,h_E) = \frac{1}{2}\left[\sqrt{\left(\frac{1}{h_E}-\frac{1}{h_M}\right)^2+\frac{4}{\lambda}\left(\frac{1}{h_E}-\frac{1}{h_M}\right)}-\left(\frac{1}{h_M}+\frac{1}{h_E}\right)\right]^+, \tag{7}$$

where $[x]^+ = \max\{0, x\}$, and the parameter λ is a constant that satisfies the power constraint in (5) with equality. The secrecy capacity is then determined by substituting this optimal power allocation policy for $P(h_M, h_E)$ in (4).

4 Only Main Channel CSI at the Transmitter

In this section, we assume that at the beginning of each coherence interval, the transmitter only knows the CSI of the main channel (legitimate receiver).

4.1 Optimal Power Allocation

We first characterize the secrecy capacity under this scenario in the following theorem.

Theorem 2 *When only the channel gain of the legitimate receiver is known at the transmitter, the secrecy capacity is given by*

$$C_s^{(M)} \;=\; \max_{P(h_M)} \iint \left[\log\left(1+h_M P(h_M)\right) - \log\left(1+h_E P(h_M)\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \;, \tag{8}$$

$$\textit{such that} \qquad \mathbb{E}\{P(h_M)\} \;\le\; \bar{P}. \tag{9}$$

Proof: A detailed proof of achievability and the converse part is provided in the Appendix. Here, we outline the scheme used to show achievability. We use the following **variable rate** transmission scheme. During a coherence interval with main channel fading state h_M, the transmitter transmits codewords at rate $\log(1 + h_M P(h_M))$ with power $P(h_M)$. This variable rate scheme relies on the assumption of large coherence intervals and ensures that when $h_E > h_M$, the mutual information between the source and the eavesdropper is upper bounded by $\log(1+h_M P(h_M))$. When $h_E \le h_M$, this mutual information will be $\log(1+h_E P(h_M))$. Averaging over all the fading states, the average rate of the main channel is given by

$$\iint \log\left(1 + h_M P(h_M)\right) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E,$$

while the information accumulated at the eavesdropper is

$$\iint \log\left(1 + \min\{h_M, h_E\} P(h_M)\right) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E.$$

Hence for a given power control policy $P(h_M)$, the achievable perfect secrecy rate is given by

$$R_s^{(M)} = \iint \left[\log\left(1 + h_M P(h_M)\right) - \log\left(1 + h_E P(h_M)\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E. \quad (10)$$

One can then optimize over all feasible power control policies $P(h_M)$ to maximize the perfect secrecy rate. Finally, we observe that our secure message is **hidden** across different fading states (please refer to our proof for more details). □

We now derive the optimal power allocation policy that achieves the secrecy capacity under the main channel CSI assumption. Similar to Theorem 1, the objective function under this case is also concave, and using the Lagrangian maximization approach for solving (8), we get the following optimality condition.

$$\frac{\partial R_s^{(M)}}{\partial P(h_M)} = \frac{h_M \Pr\left(h_E \leq h_M\right)}{1 + h_M P(h_M)} - \int_0^{h_M} \left(\frac{h_E}{1 + h_E P(h_M)}\right) f(h_E) \mathrm{d}h_E - \lambda = 0,$$

where λ is a constant that satisfies the power constraint in (9) with equality. For any main channel fading state h_M, the optimal transmit power level $P(h_M)$ is determined from the above equation. If the obtained power level turns out to be negative, then the optimal value of $P(h_M)$ is equal to 0. This follows from the concavity of the objective function in (8) w.r.t. $P(h_M)$. The solution to this optimization problem depends on the distributions $f(h_M)$ and $f(h_E)$. In the following, we focus on the Rayleigh fading scenario with $\mathbb{E}\{h_M\} = \overline{\gamma}_M$ and $\mathbb{E}\{h_E\} = \overline{\gamma}_E$ in detail. With Rayleigh fading, the objective function in (8) simplifies to

$$\begin{aligned} C_s^{(M)} &= \max_{P(h_M)} \int_0^\infty \Bigg[\left(1 - e^{-(h_M/\overline{\gamma}_E)}\right) \log\left(1 + h_M P(h_M)\right) - \\ &\qquad \int_0^{h_M} \log\left(1 + h_E P(h_M)\right) \frac{1}{\overline{\gamma}_E} e^{-(h_E/\overline{\gamma}_E)} \mathrm{d}h_E \Bigg] \frac{1}{\overline{\gamma}_M} e^{-(h_M/\overline{\gamma}_M)} \mathrm{d}h_M \\ &= \max_{P(h_M)} \int_0^\infty \Bigg[\log\left(1 + h_M P(h_M)\right) - \exp\left(\frac{1}{\overline{\gamma}_E P(h_M)}\right) \Bigg(\mathrm{Ei}\left(\frac{1}{\overline{\gamma}_E P(h_M)}\right) - \\ &\qquad \mathrm{Ei}\left(\frac{h_M}{\overline{\gamma}_E} + \frac{1}{\overline{\gamma}_E P(h_M)}\right)\Bigg)\Bigg] \frac{1}{\overline{\gamma}_M} e^{-(h_M/\overline{\gamma}_M)} \mathrm{d}h_M, \quad (11) \end{aligned}$$

where

$$\mathrm{Ei}(x) = \int_x^\infty \frac{e^{-t}}{t} \, \mathrm{d}t \, .$$

Specializing the optimality conditions to the Rayleigh fading scenario, it can be shown that the power level of the transmitter at any fading state h_M is obtained by solving the equation

$$\begin{aligned}\left(1-e^{-(h_M/\overline{\gamma}_E)}\right)\left(\frac{h_M}{1+h_M P(h_M)}\right)-\lambda-\frac{\left(1-e^{-(h_M/\overline{\gamma}_E)}\right)}{P(h_M)}+\\ \frac{\exp\left(\frac{1}{\overline{\gamma}_E P(h_M)}\right)}{\overline{\gamma}_E (P(h_M))^2}\left[\mathrm{Ei}\left(\frac{1}{\overline{\gamma}_E P(h_M)}\right)-\mathrm{Ei}\left(\frac{h_M}{\overline{\gamma}_E}+\frac{1}{\overline{\gamma}_E P(h_M)}\right)\right] &= 0.\end{aligned}$$

If there is no positive solution to this equation for a particular h_M, then we set $P(h_M)=0$. The secrecy capacity is then determined by substituting this optimal power allocation policy for $P(h_M)$ in (11).

We observe that, unlike the traditional ergodic fading scenario, achieving the optimal performance under a security constraint relies heavily on using a variable rate transmission strategy. This can be seen by evaluating the performance of a constant rate strategy where a single codeword is interleaved across infinitely many fading realizations. This interleaving will result in the eavesdropper **gaining more information**, than the destination, when its channel is better than the main channel, thereby yielding a perfect secrecy rate that is strictly smaller than that in (10). It is easy to see that the achievable perfect secrecy rate of the constant rate scheme, assuming a Gaussian codebook, is given by

$$\max_{P(h_M)}\iint\left[\log\left(1+h_M P(h_M)\right)-\log\left(1+h_E P(h_M)\right)\right]f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E\ ,$$

$$\text{such that}\qquad \mathbb{E}\{P(h_M)\}\leq\bar{P}.$$

Unlike the two previous optimization problems, the objective function in this optimization problem is not a concave function of $P(h_M)$. Using the Lagrangian formulation, we only get the following *necessary* KKT conditions for the optimal point.

$$\begin{aligned}P(h_M)\left[\lambda-\frac{h_M}{1+h_M P(h_M)}+\int\left(\frac{h_E}{1+h_E P(h_M)}\right)f(h_E)\mathrm{d}h_E\right] &= 0,\\ \lambda &\geq \frac{h_M}{1+h_M P(h_M)}-\int\left(\frac{h_E}{1+h_E P(h_M)}\right)f(h_E)\mathrm{d}h_E,\\ \mathbb{E}\{P(h_M)\} &= \bar{P}.\end{aligned}\tag{12}$$

4.2 On/Off Power Control

We now propose a transmission policy wherein the transmitter sends information only when the channel gain of the legitimate receiver h_M exceeds a pre-determined constant threshold $\tau>0$. Moreover, when $h_M>\tau$, the transmitter always uses the same power level P. However, it is crucial to adapt the rate of transmission instantaneously as $\log(1+Ph_M)$ with h_M. It is clear that for an average power constraint $\bar{P}$, the constant power level used for transmission will be

$$P=\frac{\bar{P}}{\Pr(h_M>\tau)}\ .$$

Using a similar argument as in the achievable part of Theorem 2, we get the perfect secrecy rate achieved by the proposed scheme, using Gaussian inputs, as

$$R_s^{(CP)} = \int_0^\infty \int_\tau^\infty \left[\log\left(1+h_M P\right) - \log\left(1+h_E P\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \ .$$

Specializing to the Rayleigh fading scenario, we get

$$P = \frac{\bar{P}}{\Pr(h_M > \tau)} = \bar{P} e^{(\tau/\overline{\gamma}_M)} \ ,$$

and the secrecy capacity simplifies to

$$R_s^{(CP)} = \int_\tau^\infty \int_0^{h_M} \left[\log\left(1 + h_M \bar{P} e^{(\tau/\overline{\gamma}_M)}\right) - \log\left(1 + h_E \bar{P} e^{(\tau/\overline{\gamma}_M)}\right)\right] \frac{1}{\overline{\gamma}_M} e^{-(h_M/\overline{\gamma}_M)} \frac{1}{\overline{\gamma}_E} e^{-(h_E/\overline{\gamma}_E)} \mathrm{d}h_E \mathrm{d}h_M \ ,$$

which then simplifies to

$$\begin{aligned} R_s^{(CP)} &= e^{-(\tau/\overline{\gamma}_M)} \log\left(1 + \tau \bar{P} e^{(\tau/\overline{\gamma}_M)}\right) + \exp\left(\frac{1}{\overline{\gamma}_M \bar{P} e^{(\tau/\overline{\gamma}_M)}}\right) \mathrm{Ei}\left(\frac{\tau}{\overline{\gamma}_M} + \frac{1}{\overline{\gamma}_M \bar{P} e^{(\tau/\overline{\gamma}_M)}}\right) \\ &+ \exp\left(\frac{1}{\overline{\gamma}_E \bar{P} e^{(\tau/\overline{\gamma}_M)}} - \frac{\tau}{\overline{\gamma}_M}\right) \left[\mathrm{Ei}\left(\frac{\tau}{\overline{\gamma}_E} + \frac{1}{\overline{\gamma}_E \bar{P} e^{(\tau/\overline{\gamma}_M)}}\right) - \mathrm{Ei}\left(\frac{1}{\overline{\gamma}_E \bar{P} e^{(\tau/\overline{\gamma}_M)}}\right)\right] \\ &- \exp\left(\frac{\left[\frac{1}{\overline{\gamma}_M} + \frac{1}{\overline{\gamma}_E}\right]}{\bar{P} e^{(\tau/\overline{\gamma}_M)}}\right) \mathrm{Ei}\left(\left[\frac{1}{\overline{\gamma}_M} + \frac{1}{\overline{\gamma}_E}\right]\left[\tau + \frac{1}{\bar{P} e^{(\tau/\overline{\gamma}_M)}}\right]\right) \ . \end{aligned}$$

One can then optimize over the threshold τ to get the maximum achievable perfect secrecy rate.

Finally, we establish the asymptotic optimality of this on/off scheme as the available average transmission power $\bar{P} \to \infty$. For the on/off power allocation policy, we have

$$R_s^{(CP)} = \lim_{\bar{P}\to\infty} \int_{\tau^*}^\infty \int_0^{h_M} \log\left(\frac{1 + h_M P}{1 + h_E P}\right) f(h_M) f(h_E) \mathrm{d}h_E \mathrm{d}h_M.$$

Taking $\tau^* = 0$, we get $P = \bar{P}$ and

$$\begin{aligned} R_s^{(CP)} &\geq \lim_{\bar{P}\to\infty} \int_0^\infty \int_0^{h_M} \log\left(\frac{(1/\bar{P}) + h_M}{(1/\bar{P}) + h_E}\right) f(h_M) f(h_E) \mathrm{d}h_E \mathrm{d}h_M \\ &\stackrel{(a)}{=} \int_0^\infty \int_0^{h_M} \lim_{\bar{P}\to\infty} \log\left(\frac{(1/\bar{P}) + h_M}{(1/\bar{P}) + h_E}\right) f(h_M) f(h_E) \mathrm{d}h_E \mathrm{d}h_M \\ &= \int_0^\infty \int_0^{h_M} \log\left(\frac{h_M}{h_E}\right) f(h_M) f(h_E) \mathrm{d}h_E \mathrm{d}h_M = \mathbb{E}_{\{h_M > h_E\}} \left\{\log\left(\frac{h_M}{h_E}\right)\right\}, \end{aligned} \quad (13)$$

where (a) follows from the Dominated Convergence Theorem, since

$$\left|\log\left(\frac{(1/\bar{P}) + h_M}{(1/\bar{P}) + h_E}\right)\right| \leq \left|\log\left(\frac{h_M}{h_E}\right)\right|, \qquad \forall \bar{P} \quad \text{when } h_M > h_E,$$

$$\text{and} \qquad \int_0^\infty \int_0^{h_M} \log\left(\frac{h_M}{h_E}\right) f(h_M) f(h_E) \mathrm{d}h_E \mathrm{d}h_M < \infty,$$

since $\mathbb{E}\{h_M\} < \infty$, $\left| \int_0^1 \log x \, \mathrm{d}x \right| = 1 < \infty$ and $f(h_M), f(h_E)$ are continuous and bounded.

Now under the full CSI assumption, we have

$$C_s^{(F)} = \mathbb{E}_{\{h_M > h_E\}} \left\{ \log\left(\frac{\frac{1}{P(h_M,h_E)} + h_M}{\frac{1}{P(h_M,h_E)} + h_E} \right) \right\} \leq \mathbb{E}_{\{h_M > h_E\}} \left\{ \log\left(\frac{h_M}{h_E}\right) \right\}. \quad (14)$$

From (13) and (14), it is clear that the proposed on/off power allocation policy that uses only the main channel CSI achieves the secrecy capacity under the full CSI assumption as $\bar{P} \to \infty$. Thus the absence of eavesdropper CSI at the transmitter does not reduce the secrecy capacity at high SNR values.

5 Numerical Results

As an additional benchmark, we first obtain the performance when the transmitter does not have any knowledge of both the main and eavesdropper channels (only receiver CSI). In this scenario, the transmitter is unable to exploit rate/power adaptation and always transmits with power $\bar{P}$. It is straightforward to see that the achievable perfect secrecy rate in this scenario (using Gaussian inputs) is given by

$$\begin{aligned} R_s^{(R)} &= \left[\int_0^\infty \int_0^\infty \left[\log\left(1 + h_M \bar{P}\right) - \log\left(1 + h_E \bar{P}\right) \right] f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \right]^+ \\ &= \left[\int_0^\infty \log\left(1 + h_M \bar{P}\right) f(h_M) \mathrm{d}h_M - \int_0^\infty \log\left(1 + h_E \bar{P}\right) f(h_E) \mathrm{d}h_E \right]^+, \end{aligned}$$

which reduces to the following for the Rayleigh fading scenario

$$R_s^{(R)} = \left[\exp\left(\frac{1}{\overline{\gamma}_M \bar{P}}\right) \mathrm{Ei}\left(\frac{1}{\overline{\gamma}_M \bar{P}}\right) - \exp\left(\frac{1}{\overline{\gamma}_E \bar{P}}\right) \mathrm{Ei}\left(\frac{1}{\overline{\gamma}_E \bar{P}}\right) \right]^+.$$

Thus when $\overline{\gamma}_E \geq \overline{\gamma}_M$, $R_s^{(R)} = 0$. The results for the Rayleigh normalized-symmetric case ($\overline{\gamma}_M = \overline{\gamma}_E = 1$) are presented in Fig. 2. It is clear that the performance of the on/off power control scheme is very close to the secrecy capacity (with only main channel CSI) for a wide range of SNRs and, as expected, approaches the secrecy capacities, under both the full CSI and main channel CSI assumptions, at high values of SNR. The performance of the constant rate scheme is much worse than the other schemes that employ rate adaptation. Here we note that the performance curve for the constant rate scheme might be a lower bound to the secrecy capacity (since the KKT conditions are necessary but not sufficient for non-convex optimization). We then consider an asymmetric scenario, wherein the eavesdropper channel is more capable than the main channel, with $\overline{\gamma}_M = 1$ and $\overline{\gamma}_E = 2$. The performance results for this scenario are plotted in Fig. 3. Again it is clear from the plot that the performance of the on/off power control scheme is optimal at high values of SNR, and that rate adaptation schemes yield higher perfect secrecy rates than constant rate transmission schemes.

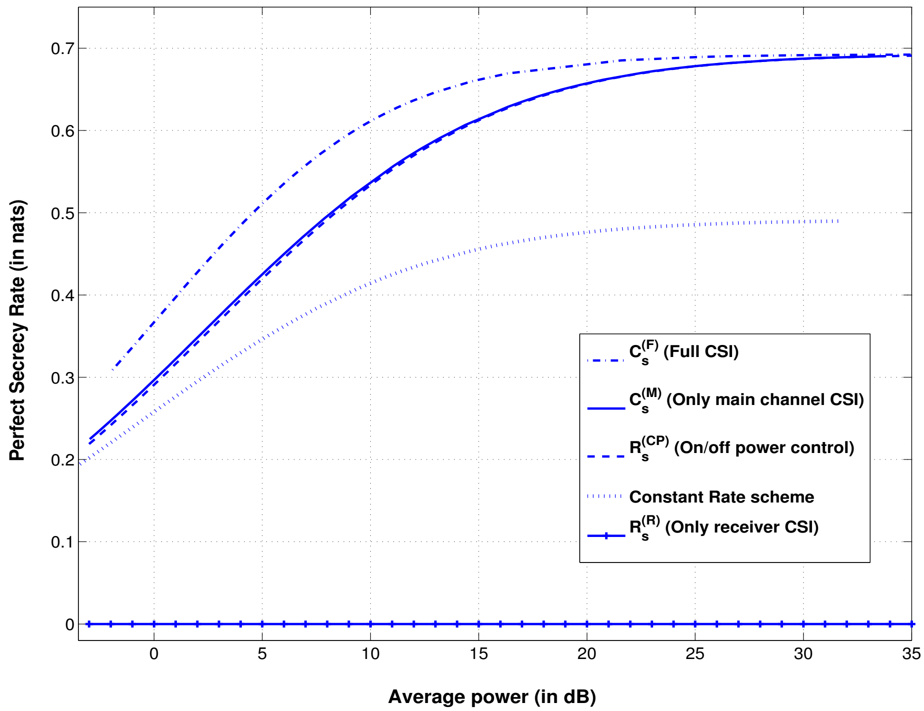

Figure 2: Performance comparison for the symmetric scenario $\overline{\gamma}_M = \overline{\gamma}_E = 1$.

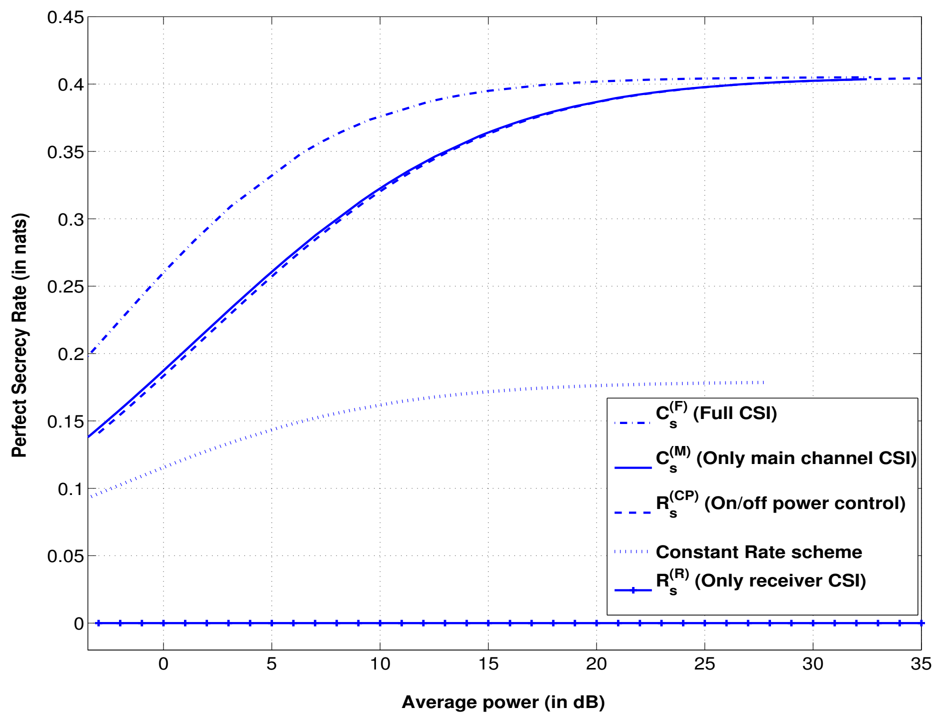

Figure 3: Performance comparison for the asymmetric scenario $\overline{\gamma}_M = 1$ and $\overline{\gamma}_E = 2$.

6 Conclusions

We have characterized the secrecy capacity of the slow fading channel with an eavesdropper under different assumptions on the available transmitter CSI. Our work establishes the interesting result that a non-zero perfectly secure rate is achievable in the fading channel even when the eavesdropper is more capable than the legitimate receiver (on the average). By contrasting this conclusion with the traditional AWGN scenario, one can see the positive impact of fading on **enhancing** the secrecy capacity. Furthermore, we proposed a low-complexity on/off power transmission scheme and established its asymptotic optimality. This optimality shows that the presence of eavesdropper CSI at the transmitter does not offer additional gains in the secrecy capacity for slow fading channels, at high enough SNR levels. The knowledge of the main channel CSI, however, is crucial since it is easy to see that the absence of this information leads to a zero secrecy capacity when the eavesdropper is more capable than the legitimate receiver on the average. Finally, our theoretical and numerical results established the critical role of appropriate rate adaptation in facilitating secure communications over slow fading channels.

A Proof of Theorem 1

We first prove the achievability of (4) by showing that for any perfect secrecy rate $R_s < C_s^{(F)}$, there exists a sequence of $(2^{nR_s}, n)$ block codes with average power $\bar{P}$, equivocation rate $R_e > R_s - \epsilon$, and probability of error $P_e^n \to 0$ as $n \to \infty$. Let $R_s = C_s^{(F)} - 3\delta$ for some $\delta > 0$. We quantize the main channel gains $h_M \in [0, M_1]$ into uniform bins $\{h_{M,i}\}_{i=1}^{q_1}$, and the eavesdropper channel gains $h_E \in [0, M_2]$ into uniform bins $\{h_{E,j}\}_{j=1}^{q_2}$. The channels are said to be in state s_{ij} ($i \in [1, q_1]$, $j \in [1, q_2]$), if $h_{M,i} \le h_M < h_{M,(i+1)}$ and $h_{E,j} \le h_E < h_{E,(j+1)}$, where $h_{M,(q_1+1)} = M_1, h_{E,(q_2+1)} = M_2$. We also define a power control policy for any state s_{ij} by

$$P(h_{M,i}, h_{E,j}) \;=\; \inf_{h_{M,i} \le h_M < h_{M,(i+1)}, h_{E,j} \le h_E < h_{E,(j+1)}} P(h_M, h_E) \;, \tag{15}$$

where $P(h_M, h_E)$ is the optimal power allocation policy in (7) that satisfies $P(h_M, h_E) = 0$ for all $h_M \le h_E$, and the power constraint

$$\int_0^\infty \int_{h_E}^\infty P(h_M, h_E) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \;\le\; \bar{P}. \tag{16}$$

Consider a time-invariant AWGN channel with channel gains $h_M \in [h_{M,i}, h_{M,(i+1)})$ and $h_E \in [h_{E,j}, h_{E,(j+1)})$. It is shown in [5,6] that for this channel, we can develop a sequence of $(2^{n_{ij}(R_s)_{ij}}, n_{ij})$ codes with codeword rate $\log(1 + h_{M,i} P(h_{M,i}, h_{E,j}))$ and perfect secrecy rate

$$(R_s)_{ij} \;=\; \Big[\log\left(1 + h_{M,i} P(h_{M,i}, h_{E,j})\right) - \log\left(1 + h_{E,(j+1)} P(h_{M,i}, h_{E,j})\right) \Big]^+ \;, \tag{17}$$

such that the average power is $P(h_{M,i}, h_{E,j})$ and with error probability $P_e^{ij} \to 0$ as $n_{ij} \to \infty$, where

$$n_{ij} = n \;\Pr\left(h_{M,i} \le h_M < h_{M,(i+1)}, h_{E,j} \le h_E < h_{E,(j+1)}\right)$$

for sufficiently large n. Note that the expression in (17) is obtained by considering the worst case scenario $h_M = h_{M,i}, h_E = h_{E,(j+1)}$ that yields the smallest perfect secrecy rate.

For transmitting the message index $w \in \{1, \cdots, 2^{nR_s}\}$, we first map w to the indices $\{w_{ij}\}$ by dividing the nR_s bits which determine the message index into sets of $n_{ij}(R_s)_{ij}$ bits. The transmitter uses a multiplexing strategy and transmits codewords $\{x_{w_{ij}}\}$ at codeword rate $\log(1 + h_{M,i}P(h_{M,i}, h_{E,j}))$ and perfect secrecy rate $(R_s)_{ij}$, when the channel is in state s_{ij}. As $n \to \infty$, this scheme achieves the perfect secrecy rate (using the ergodicity of the channel),

$$R_s = \sum_{i=1}^{q_1}\sum_{j=1}^{q_2}\left[\log\left(\frac{1 + h_{M,i}P(h_{M,i}, h_{E,j})}{1 + h_{E,(j+1)}P(h_{M,i}, h_{E,j})}\right)\right]^+ \Pr\left(h_{M,i} \le h_M < h_{M,(i+1)}, h_{E,j} \le h_E < h_{E,(j+1)}\right).$$

Thus for a fixed δ, we can find a sufficiently large n such that

$$R_s \ge \sum_{i=1}^{q_1}\sum_{j=1}^{q_2}\left[\log\left(\frac{1 + h_{M,i}P(h_{M,i}, h_{E,j})}{1 + h_{E,(j+1)}P(h_{M,i}, h_{E,j})}\right)\right]^+ \Pr\left(h_{M,i} \le h_M < h_{M,(i+1)}, h_{E,j} \le h_E < h_{E,(j+1)}\right) - \delta. \tag{18}$$

For asymptotically large n, using the ergodicity of the channel, the average power of the multiplexing scheme satisfies

$$\begin{aligned}&\sum_{i=1}^{q_1}\sum_{j=1}^{q_2} P(h_{M,i}, h_{E,j}) \int_{h_{M,i}}^{h_{M,(i+1)}} \int_{h_{E,j}}^{h_{E,(j+1)}} f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E \\ &\overset{(a)}{\le} \int_0^\infty \int_0^\infty P(h_M, h_E)f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E \overset{(b)}{\le} \bar{P},\end{aligned}$$

where (a) follows from the definition of $P(h_{M,i}, h_{E,j})$ in (15) and (b) follows from (16). Moreover, the error probability of the multiplexing scheme is upper bounded by

$$P_e^n \le \sum_{i=1}^{q_1}\sum_{j=1}^{q_2} P_e^{ij} \to 0, \quad \text{as } n \to \infty.$$

Now since

$$\begin{aligned} C_s^{(F)} &= \int_0^\infty \int_0^\infty [\log(1 + h_M P(h_M, h_E)) - \log(1 + h_E P(h_M, h_E))]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E \\ &\le \int_0^\infty \int_{h_E}^\infty \log\left(\frac{h_M}{h_E}\right) f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E < \infty, \end{aligned}$$

(because $\mathbb{E}\{h_M\} < \infty$, $\left|\int_0^1 \log x \, \mathrm{d}x\right| = 1 < \infty$ and $f(h_M), f(h_E)$ are continuous and bounded), there exist M_1 and M_2 for a fixed δ such that

$$\begin{aligned} \int_0^{M_1} \int_{M_2}^\infty [\log(1 + h_M P(h_M, h_E)) - \log(1 + h_E P(h_M, h_E))]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E &< \frac{\delta}{3}, \\ \int_{M_1}^\infty \int_0^{M_2} [\log(1 + h_M P(h_M, h_E)) - \log(1 + h_E P(h_M, h_E))]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E &< \frac{\delta}{3}, \\ \int_{M_1}^\infty \int_{M_2}^\infty [\log(1 + h_M P(h_M, h_E)) - \log(1 + h_E P(h_M, h_E))]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E &< \frac{\delta}{3}. \end{aligned} \tag{19}$$

Moreover, for fixed M_1 and M_2, the dominated convergence theorem implies that

$$\lim_{(q_1,q_2)\to\infty} \sum_{i=1}^{q_1}\sum_{j=1}^{q_2}\left[\log\left(\frac{1+h_{M,i}P(h_{M,i},h_{E,j})}{1+h_{E,(j+1)}P(h_{M,i},h_{E,j})}\right)\right]^+ \Pr\left(h_{M,i}\le h_M < h_{M,(i+1)}, h_{E,j}\le h_E < h_{E,(j+1)}\right)$$

$$= \lim_{(q_1,q_2)\to\infty} \sum_{i=1}^{q_1}\sum_{j=1}^{q_2}\int_{h_{M,i}}^{h_{M,(i+1)}}\int_{h_{E,j}}^{h_{E,(j+1)}}\left[\log\left(\frac{1+h_{M,i}P(h_{M,i},h_{E,j})}{1+h_{E,(j+1)}P(h_{M,i},h_{E,j})}\right)\right]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E$$

$$= \int_0^{M_1}\int_0^{M_2}\left[\log\left(\frac{1+h_MP(h_M,h_E)}{1+h_EP(h_M,h_E)}\right)\right]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E. \quad (20)$$

Choosing M_1, M_2 that satisfy (19) and combining (19) and (20), we see that for a given δ, there exist sufficiently large q_1, q_2 such that

$$\sum_{i=1}^{q_1}\sum_{j=1}^{q_2}\left[\log\left(\frac{1+h_{M,i}P(h_{M,i},h_{E,j})}{1+h_{E,(j+1)}P(h_{M,i},h_{E,j})}\right)\right]^+ \Pr\left(h_{M,i}\le h_M < h_{M,(i+1)}, h_{E,j}\le h_E < h_{E,(j+1)}\right)$$

$$\ge \int_0^{\infty}\int_0^{\infty}\left[\log\left(\frac{1+h_MP(h_M,h_E)}{1+h_EP(h_M,h_E)}\right)\right]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E \;-\; 2\delta. \quad (21)$$

Combining (18) and (21), we get the desired result.

We now prove the converse part by showing that for any perfect secrecy rate R_s with equivocation rate $R_e > R_s - \epsilon$ and error probability $P_e^n \to 0$ as $n \to \infty$, there exists a power allocation policy $P(h_M, h_E)$ satisfying the average power constraint, such that

$$R_s \;\le\; \iint [\log(1+h_MP(h_M,h_E)) - \log(1+h_EP(h_M,h_E))]^+ f(h_M)f(h_E)\mathrm{d}h_M\mathrm{d}h_E.$$

Consider any sequence of $(2^{nR_s}, n)$ codes with perfect secrecy rate R_s and equivocation rate R_e, such that $R_e > R_s - \epsilon$, with average power less than or equal to $\bar{P}$ and error probability $P_e^n \to 0$ as $n \to \infty$. Let $N(h_M, h_E)$ denote the number of times the channel is in fading state (h_M, h_E) over the interval $[0, n]$. Also let $P^n(h_M, h_E) = \mathbb{E}\big\{\sum_{i=1}^n |x_w(i)|^2 \mathbf{1}_{\{h_M(i)=h_M, h_E(i)=h_E\}}\big\}$, where $\{x_w\}$ are the codewords corresponding to the message w and the expectation is taken over all codewords. We note that the equivocation $H(W|Z^n, h_M^n, h_E^n)$ only depends on the marginal distribution of Z^n, and thus does not depend on whether $Z(i)$ is a physically or stochastically degraded version of $Y(i)$ or vice versa. Hence we assume in the following derivation that for any fading state, either $Z(i)$ is a physically degraded version of $Y(i)$ or vice versa (since the noise processes are Gaussian), depending on the instantaneous channel state. Thus we have

$$\begin{aligned}
nR_e &= H(W|Z^n, h_M^n, h_E^n)\\
&\overset{(a)}{\le} H(W|Z^n, h_M^n, h_E^n) - H(W|Z^n, Y^n, h_M^n, h_E^n) + n\delta_n\\
&= I(W; Y^n|Z^n, h_M^n, h_E^n) + n\delta_n\\
&\overset{(b)}{\le} I(X^n; Y^n|Z^n, h_M^n, h_E^n) + n\delta_n\\
&= H(Y^n|Z^n, h_M^n, h_E^n) - H(Y^n|X^n, Z^n, h_M^n, h_E^n) + n\delta_n\\
&= \sum_{i=1}^n \left[H(Y(i)|Y^{i-1}, Z^n, h_M^n, h_E^n) - H(Y(i)|Y^{i-1}, X^n, Z^n, h_M^n, h_E^n)\right] + n\delta_n
\end{aligned}$$

$$
\begin{aligned}
&\overset{(c)}{\leq} \sum_{i=1}^{n} [H(Y(i)|Z(i), h_M(i), h_E(i)) - H(Y(i)|X(i), Z(i), h_M(i), h_E(i))] + n\delta_n \\
&= \sum_{i=1}^{n} I(X(i); Y(i)|Z(i), h_M(i), h_E(i)) + n\delta_n \\
&= \sum_{i=1}^{n} \iint I(X; Y|Z, h_M, h_E) \mathbf{1}_{\{h_M(i)=h_M, h_E(i)=h_E\}} \mathrm{d}h_M \mathrm{d}h_E + n\delta_n \qquad (22) \\
&= \iint I(X; Y|Z, h_M, h_E) N(h_M, h_E) \mathrm{d}h_M \mathrm{d}h_E + n\delta_n \\
&\overset{(d)}{\leq} \iint N(h_M, h_E) \left[\log(1 + h_M P^n(h_M, h_E)) - \log(1 + h_E P^n(h_M, h_E))\right]^+ \mathrm{d}h_M \mathrm{d}h_E + n\delta_n.
\end{aligned}
$$

In the above derivation, (a) follows from the Fano inequality, (b) follows from the data processing inequality since $W \rightarrow X^n \rightarrow (Y^n, Z^n)$ forms a Markov chain, (c) follows from the fact that conditioning reduces entropy and from the memoryless property of the channel, (d) follows from the fact that given h_M and h_E, the fading channel reduces to an AWGN channel with channel gains (h_M, h_E) and average transmission power $P^n(h_M, h_E)$, for which

$$
I(X; Y|Z, h_M, h_E) \;\leq\; \left[\log(1 + h_M P^n(h_M, h_E)) - \log(1 + h_E P^n(h_M, h_E))\right]^+ \;,
$$

as shown in [5, 6]. Since the codewords satisfy the power constraint, we have

$$
\iint P^n(h_M, h_E) \left(\frac{N(h_M, h_E)}{n}\right) \mathrm{d}h_M \mathrm{d}h_E \;\leq\; \bar{P}.
$$

For any h_M, h_E such that $f(h_M, h_E) \neq 0$, $\{P^n(h_M, h_E)\}$ are bounded sequences in n. Thus there exists a subsequence that converges to a limit $P(h_M, h_E)$ as $n \rightarrow \infty$. Since for each n, the power constraint is satisfied, we have

$$
\iint P(h_M, h_E) f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E \;\leq\; \bar{P}. \qquad (23)
$$

Now, we have

$$
R_e \;\leq\; \iint \frac{N(h_M, h_E)}{n} \left[\log\left(\frac{1 + h_M P^n(h_M, h_E)}{1 + h_E P^n(h_M, h_E)}\right)\right]^+ \mathrm{d}h_M \mathrm{d}h_E + \delta_n.
$$

Taking the limit along the convergent subsequence and using the ergodicity of the channel, we get

$$
R_e \;\leq\; \iint \left[\log\left(\frac{1 + h_M P(h_M, h_E)}{1 + h_E P(h_M, h_E)}\right)\right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E + \delta_n.
$$

The claim is thus proved.

B Proof of Theorem 2

Let $R_s = C_s^{(M)} - \delta$ for some small $\delta > 0$. Let $n = n_1 m$, where n_1 represents the number of symbols transmitted in each coherence interval, and m represents the number of coherence intervals over which the message W is transmitted. Let $R = \mathbb{E}\{\log(1 + h_M P(h_M))\} - \epsilon$. We first generate all binary sequences $\{\mathbf{V}\}$ of length nR and then independently assign each of them randomly to one of 2^{nR_s} groups, according to a uniform distribution. This ensures that any of the sequences are equally likely to be within any of the groups. Each secret message $w \in \{1, \cdots, 2^{nR_s}\}$ is then assigned a group $\mathbf{V}(w)$. To encode a particular message w, the stochastic encoder randomly selects a sequence $\mathbf{v}$ from the corresponding group $\mathbf{V}(w)$, according to a uniform distribution. This sequence $\mathbf{v}$ consisting of nR bits is then sub-divided into independent blocks $\{\mathbf{v}(1), \cdots, \mathbf{v}(m)\}$, where the block $\mathbf{v}(i)$ consists of $n_1 [\log(1 + h_M(i)P(h_M(i))) - \epsilon]$ bits, and is transmitted in the i^{th} coherence interval ($i \in \{1, \cdots, m\}$). As $m \to \infty$, using the ergodicity of the channel, we have

$$\lim_{m\to\infty} \sum_{i=1}^{m} n_1 [\log(1 + h_M(i)P(h_M(i))) - \epsilon] = n_1 m [\mathbb{E}\{\log(1 + h_M P(h_M))\} - \epsilon] = nR.$$

We then generate i.i.d. Gaussian codebooks $\{X^{n_1}(i) : i = 1, \cdots, m\}$ consisting of $2^{n_1[\log(1+h_M(i)P(h_M(i)))-\epsilon]}$ codewords, each of length n_1 symbols. In the i^{th} coherence interval, the transmitter encodes the block $\mathbf{v}(i)$ into the codeword $x^{n_1}(i)$, which is then transmitted over the fading channel. The legitimate receiver receives $y^{n_1}(i)$ while the eavesdropper receives $z^{n_1}(i)$ in the i^{th} coherence interval. The equivocation rate at the eavesdropper can then be lower bounded as follows.

$$\begin{aligned}
nR_e &= H(W|Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&= H(W, Z^{n_1}(1), \cdots, Z^{n_1}(m)|h_M^n, h_E^n) - H(Z^{n_1}(1), \cdots, Z^{n_1}(m)|h_M^n, h_E^n) \\
&= H(W, Z^{n_1}(1), \cdots, Z^{n_1}(m), X^{n_1}(1), \cdots, X^{n_1}(m)|h_M^n, h_E^n) - H(Z^{n_1}(1), \cdots, Z^{n_1}(m)|h_M^n, h_E^n) \\
&\qquad - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&= H(X^{n_1}(1), \cdots, X^{n_1}(m)|h_M^n, h_E^n) + H(W, Z^{n_1}(1), \cdots, Z^{n_1}(m)|X^{n_1}(1), \cdots, X^{n_1}(m), h_M^n, h_E^n) \\
&\quad - H(Z^{n_1}(1), \cdots, Z^{n_1}(m)|h_M^n, h_E^n) - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&\geq H(X^{n_1}(1), \cdots, X^{n_1}(m)|h_M^n, h_E^n) + H(Z^{n_1}(1), \cdots, Z^{n_1}(m)|X^{n_1}(1), \cdots, X^{n_1}(m), h_M^n, h_E^n) \\
&\quad - H(Z^{n_1}(1), \cdots, Z^{n_1}(m)|h_M^n, h_E^n) - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&= H(X^{n_1}(1), \cdots, X^{n_1}(m)|h_M^n, h_E^n) - I(Z^{n_1}(1), \cdots, Z^{n_1}(m); X^{n_1}(1), \cdots, X^{n_1}(m)|h_M^n, h_E^n) \\
&\qquad - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&= H(X^{n_1}(1), \cdots, X^{n_1}(m)|Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&\qquad - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&\overset{(a)}{=} \sum_{i=1}^{m} H(X^{n_1}(i)|Z^{n_1}(i), h_M(i), h_E(i)) - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&\overset{(b)}{\geq} \sum_{i\in\mathcal{N}_m} H(X^{n_1}(i)|Z^{n_1}(i), h_M(i), h_E(i)) - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&= \sum_{i\in\mathcal{N}_m} [H(X^{n_1}(i)|h_M(i), h_E(i)) - I(X^{n_1}(i); Z^{n_1}(i)|h_M(i), h_E(i))] \\
&\qquad - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \\
&\geq \sum_{i\in\mathcal{N}_m} n_1 [\log(1 + h_M(i)P(h_M(i))) - \log(1 + h_E(i)P(h_M(i))) - \epsilon] \\
&\qquad - H(X^{n_1}(1), \cdots, X^{n_1}(m)|W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n)
\end{aligned}$$

$$\begin{aligned}
&\geq \sum_{i=1}^{m} n_1 \left\{ \left[\log\left(1+h_M(i)P(h_M(i))\right) - \log\left(1+h_E(i)P(h_M(i))\right) \right]^+ - \epsilon \right\} \\
&\qquad\qquad\qquad\qquad -H(X^{n_1}(1),\cdots,X^{n_1}(m)|W,Z^{n_1}(1),\cdots,Z^{n_1}(m),h_M^n,h_E^n) \\
&\overset{(c)}{=} nC_s^{(M)} - H(X^{n_1}(1),\cdots,X^{n_1}(m)|W,Z^{n_1}(1),\cdots,Z^{n_1}(m),h_M^n,h_E^n) - n\epsilon. \qquad (24)
\end{aligned}$$

In the above derivation, (a) follows from the memoryless property of the channel and the independence of the $X^{n_1}(i)$'s, (b) is obtained by removing all those terms which correspond to the coherence intervals $i \notin \mathcal{N}_m$, where $\mathcal{N}_m = \{i \in \{1,\cdots,m\} : h_M(i) > h_E(i)\}$, and (c) follows from the ergodicity of the channel as $m \to \infty$.

Now we show that the term $H(X^{n_1}(1),\cdots,X^{n_1}(m)|W,Z^{n_1}(1),\cdots,Z^{n_1}(m),h_M^n,h_E^n)$ vanishes as $m, n_1 \to \infty$ by using a list decoding argument. In this list decoding, at coherence interval i, the eavesdropper first constructs a list $\mathcal{L}_i$ such that $x^{n_1}(i) \in \mathcal{L}_i$ if $(x^{n_1}(i), z^{n_1}(i))$ are jointly typical. Let $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2 \times \cdots \times \mathcal{L}_m$. Given w, the eavesdropper declares that $\hat{x}^n = (x^{n_1}(1),\cdots,x^{n_1}(m))$ was transmitted, if $\hat{x}^n$ is the only codeword such that $\hat{x}^n \in B(w) \cap \mathcal{L}$, where $B(w)$ is the set of codewords corresponding to the message w. If the eavesdropper finds none or more than one such sequence, then it declares an error. Hence, there are two type of error events: 1) $\mathcal{E}_1$: the transmitted codeword x_t^n is not in $\mathcal{L}$, 2) $\mathcal{E}_2$: $\exists x^n \neq x_t^n$ such that $x^n \in B(w) \cap \mathcal{L}$. Thus the error probability $\Pr(\hat{x}^n \neq x_t^n) = \Pr(\mathcal{E}_1 \cup \mathcal{E}_2) \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2)$. Based on the AEP, we know that $\Pr(\mathcal{E}_1) \leq \epsilon_1$. In order to bound $\Pr(\mathcal{E}_2)$, we first bound the size of $\mathcal{L}_i$. We let

$$\phi_i(x^{n_1}(i)|z^{n_1}(i)) = \begin{cases} 1, & \text{when } (x^{n_1}(i), z^{n_1}(i)) \text{ are jointly typical,} \\ 0, & \text{otherwise.} \end{cases} \qquad (25)$$

Now

$$\begin{aligned}
\mathbb{E}\{\|\mathcal{L}_i\|\} &= \mathbb{E}\left\{ \sum_{x^{n_1}(i)} \phi_i(x^{n_1}(i)|z^{n_1}(i)) \right\} \\
&\leq \mathbb{E}\left\{ 1 + \sum_{x^{n_1}(i) \neq x_t^{n_1}(i)} \phi_i(x^{n_1}(i)|z^{n_1}(i)) \right\} \\
&\leq 1 + \sum_{x^{n_1}(i) \neq x_t^{n_1}(i)} \mathbb{E}\left\{ \phi_i(x^{n_1}(i)|z^{n_1}(i)) \right\} \\
&\leq 1 + 2^{n_1 [\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i))) - \epsilon]} \\
&\leq 2^{n_1 \left([\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i))) - \epsilon]^+ + \frac{1}{n_1} \right)} . \qquad (26)
\end{aligned}$$

Hence

$$\mathbb{E}\{\|\mathcal{L}\|\} = \prod_{i=1}^{m} \mathbb{E}\{\|\mathcal{L}_i\|\} \leq 2^{\sum_{i=1}^{m} n_1 \left([\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i))) - \epsilon]^+ + \frac{1}{n_1} \right)} . \qquad (27)$$

Thus

$$
\begin{aligned}
\Pr(\mathcal{E}_2) &\leq \mathbb{E}\left\{ \sum_{x^n \in \mathcal{L}, x^n \neq x_t^n} \Pr(x^n \in B(w)) \right\} \\
&\stackrel{(a)}{\leq} \mathbb{E}\left\{ \|\mathcal{L}\| 2^{-nR_s} \right\} \\
&\leq 2^{-nR_s} 2^{\sum_{i=1}^{m} n_1 \left([\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i))) - \epsilon]^+ + \frac{1}{n_1} \right)} \\
&\leq 2^{-n\left(R_s - \frac{1}{m} \sum_{i=1}^{m} \left([\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i))) - \epsilon]^+ + \frac{1}{n_1} \right) \right)} \\
&= 2^{-n\left(R_s - \frac{1}{m} \sum_{i=1}^{m} \left([\log(1+h_M(i)P(h_M(i))) - \log(1+h_E(i)P(h_M(i)))]^+ + \frac{1}{n_1} \right) + \frac{|\mathcal{N}_m|\epsilon}{m} \right)},
\end{aligned}
\tag{28}
$$

where (a) follows from the uniform distribution of the codewords in $B(w)$. Now as $n_1 \to \infty$ and $m \to \infty$, we get

$$
\Pr(\mathcal{E}_2) \;\leq\; 2^{-n(C_s - \delta - C_s + c\epsilon)} \;=\; 2^{-n(c\epsilon - \delta)},
$$

where $c = \Pr(h_M > h_E)$. Thus, by choosing $\epsilon > (\delta / c)$, the error probability $\Pr(\mathcal{E}_2) \to 0$ as $n \to \infty$. Now using Fano's inequality, we get

$$
H(X^{n_1}(1), \cdots, X^{n_1}(m) | W, Z^{n_1}(1), \cdots, Z^{n_1}(m), h_M^n, h_E^n) \;\leq\; n\delta_n \qquad \to 0 \qquad \text{as } n \to \infty.
$$

Combining this with (24), we get the desired result.

For the converse part, consider any sequence of $(2^{nR_s}, n)$ codes with perfect secrecy rate R_s and equivocation rate R_e, such that $R_e > R_s - \epsilon$, with average power less than or equal to $\bar{P}$ and error probability $P_e^n \to 0$ as $n \to \infty$. We follow the same steps used in the proof of the converse in Theorem 1 with the only difference that now the transmission power $P^n(.)$ only depends on h_M. From (22), we get

$$
\begin{aligned}
nR_e &\leq \sum_{i=1}^{n} \iint I(X;Y|Z,h_M,h_E) \mathbf{1}_{\{h_M(i)=h_M, h_E(i)=h_E\}} \mathrm{d}h_M \mathrm{d}h_E + n\delta_n \\
&= \iint I(X;Y|Z,h_M,h_E) N(h_M,h_E) \mathrm{d}h_M \mathrm{d}h_E + n\delta_n \\
&\leq \iint N(h_M,h_E) \left[\log(1+h_M P^n(h_M)) - \log(1+h_E P^n(h_M)) \right]^+ \mathrm{d}h_M \mathrm{d}h_E + n\delta_n.
\end{aligned}
$$

This follows from the fact that given h_M and h_E, the fading channel reduces to an AWGN channel with channel gains (h_M, h_E) and average transmission power $P^n(h_M)$, for which Gaussian inputs are known to be optimal [5, 6].

Similar to the proof of Theorem 1, we take the limit over the convergent subsequence and use the ergodicity of the channel to obtain

$$
R_e \leq \iint \left[\log(1+h_M P(h_M)) - \log(1+h_E P(h_M)) \right]^+ f(h_M) f(h_E) \mathrm{d}h_M \mathrm{d}h_E + \delta_n, \tag{29}
$$

where $\mathbb{E}\{P(h_M)\} \leq \bar{P}$. The claim is thus proved.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656-715, Oct. 1949.

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

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. on Information Theory*, vol. 24, pp. 339-348, May 1978.

[4] G. Caire, G. Taricco and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. on Information Theory*, vol. 45, no. 5, pp. 1468-1489, July 1999.

[5] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian Wire-Tap Channel," *IEEE Trans. on Information Theory*, vol. 24, pp. 451-456, July 1978.

[6] Y. Liang and H. Vincent Poor, "Generalized multiple access channels with confidential messages," *Submitted to IEEE Trans. on Information Theory*, April 2006.

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

[8] P. Parada and R. Blahut, "Secrecy Capacity of SIMO and Slow Fading Channels," *Proc. of ISIT 2005*, pp. 2152–2155, Sep. 2005.

[9] J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels," *Proc. of ISIT 2006*, July 2006.

[10] P. K. Gopala, L. Lai and H. El Gamal, "The Secrecy Capacity of the Fading Wiretap Channel," *Submitted to ICASSP 2007*, Sep. 2006.