Fading Wiretap Channel with No CSI Anywhere

Pritam Mukherjee Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
pritamm@umd.edu *ulukus@umd.edu*

***Abstract*—We consider the fast Rayleigh fading wiretap channel, over which a legitimate transmitter wishes to have secure communication with a legitimate receiver in the presence of an eavesdropper. We consider an average power constraint on the input, and assume that no channel state information (CSI) is available to any user. We show that the input distribution that achieves the secrecy capacity for this wiretap channel is discrete with a finite number of mass points.**

I. Introduction

We consider the wiretap channel where a legitimate transmitter wishes to have information-theoretically secure communication with a legitimate receiver in the presence of an eavesdropper. The wiretap channel was introduced by Shannon [1] for the case of noiseless channels, where it was shown that secure keys and one-time-pad encryption was necessary for secure communications. The noisy wiretap channel was introduced by Wyner, who determined the capacity equivocation region for the degraded case [2]. Csiszar and Korner generalized his result to arbitrary, not necessarily degraded, wiretap channels [3]. Leung-Yan-Cheong and Hellman determined the capacity-equivocation region of the Gaussian wiretap channel [4], and showed that the optimal channel input was Gaussian.

In this paper, we consider the Gaussian wiretap channel under Rayleigh fading, where the channel gains of both the legitimate link and the eavesdropper link fade in an independent identically distributed (i.i.d.) fashion from one symbol to the next with a Rayleigh distribution. This models a fast fading wireless communication channel with coherence time of one symbol duration. The fading wiretap channel was considered under several different channel state information (CSI) availability conditions. References [5]–[8] considered the fading wiretap channel where all parties had complete and perfect CSI of both links. Modeling the fading wiretap under full CSI as a bank of independent parallel channels, these references showed that the capacity achieving channel inputs were independent Gaussian random variables in all parallel channels, and the variances of these random variables were found via water-filling. Reference [9] considered the case where the transmitter had the legitimate channel's CSI but no eavesdropper CSI under the assumption of infinite coherence times for channel fading, where the channel state of the eavesdropper, although unknown at the transmitter, remained constant for an infinite duration, and showed the optimality of Gaussian channel inputs in this model. Reference [10] considered the same model under a fast fading condition, i.e., when the eavesdropper channel gain is unknown at the transmitter and also varies at the order of symbol duration, and showed that MQAM signaling or Gaussian signaling with added Gaussian artificial noise, may outperform plain Gaussian signaling.

In this paper, we consider a fast fading Rayleigh wiretap channel where neither the transmitter nor the receivers have any CSI. Typically, the way CSI becomes available at the terminals is via the receivers measuring it and feeding it back to the transmitters. Under a fast fading condition, the channel may change too quickly for receivers to estimate it. In addition, the eavesdropper will not feed her CSI estimate back even if she measures it. For this system model, we determine the exact secrecy capacity. In particular we show that discrete channel inputs are optimum. We use the proof technique that was originally developed by Smith [11] to evaluate the channel capacity of an amplitude constrained Gaussian channel. This technique was further used and extended by Abou-Faycal et al. [12] to determine the channel capacity of a fast fading Rayleigh channel under an average power constraint. Our paper may be viewed as a wiretap version of Abou-Faycal et al.'s paper, which considered only reliable communication between two terminals, whereas we consider both reliability and secrecy. Our work is also closely related to [13] which considers secret key generation for a similar channel model.

We first show that this channel is equivalent to a degraded wiretap channel. This implies that no channel prefixing is needed [2]. We then consider the secrecy rate, which is the difference of mutual informations, as the objective function, which is concave, and determine the optimal input distribution as the result of a functional optimization problem. We obtain the KKT optimality conditions, and extend these conditions to the complex plane and reach a contradiction using the identity theorem to conclude that the optimum input distribution cannot have an infinite support over any finite interval. We then show that the optimal distribution has a finite support.

II. System Model, Definitions and Preliminaries

The fast Rayleigh fading wiretap channel is given by:

$$V_i = A_i U_i + N_{1i} \tag{1}$$

$$W_i = B_i U_i + N_{2i} \tag{2}$$

where U_i is the channel input, V_i and W_i are the channel outputs of the legitimate receiver and the eavesdropper, respectively, and A_i and B_i are identically distributed

This work was supported by NSF Grants CNS 09-64632, CCF 09-64645, CCF 10-18185 and CNS 11-47811.

complex circular Gaussian random variables with zero-mean and variance σ_h^2, representing fading. The realizations of A_i and B_i are unknown to all users, though their statistics are known. The noise terms N_{1i} and N_{2i} are zero-mean complex circular Gaussian random variables with variances σ_1^2 and σ_2^2, respectively, with $\sigma_2^2 > \sigma_1^2$. The random variables A_i, B_i, N_{1i}, N_{2i} are i.i.d. in time. The channel input is average power constrained: $\mathbb{E}\left[|U_i|^2\right] \leq P$.

As in [12], since the channel is stationary and memoryless, we can drop the time index i without any loss of generality. Also, since the phases of the fading parameters A and B are uniform, $|V|^2$ and $|W|^2$ are sufficient statistics to characterize the conditional distributions of V and W respectively, given the input U. Conditioned on $|U|$, $|V|^2$ and $|W|^2$ are exponentially distributed with parameters $\frac{1}{\sigma_h^2|u|^2+\sigma_1^2}$ and $\frac{1}{\sigma_h^2|u|^2+\sigma_2^2}$. We let $Y = |V|^2$, $Z = |W|^2$ and $X = |U|$, then

$$p_{Y|X}(y|x) = \frac{1}{\sigma_h^2 x^2 + \sigma_1^2} \exp\left[-\frac{y}{\sigma_h^2 x^2 + \sigma_1^2}\right] \quad (3)$$

$$p_{Z|X}(z|x) = \frac{1}{\sigma_h^2 x^2 + \sigma_2^2} \exp\left[-\frac{z}{\sigma_h^2 x^2 + \sigma_2^2}\right] \quad (4)$$

The transmitter sends a message M, uniformly chosen from $\mathcal{M}$, by encoding it to an n-length codeword $U^n = \varphi(M)$ using a stochastic encoding function φ. The legitimate receiver detects the message $\hat{M} = \psi(V^n)$ using a decoding function ψ. The rate of communication is $R = \frac{1}{n}\log|\mathcal{M}|$, and the probability of error is $P_e = \mathbb{P}[\hat{M} \neq M]$. The secrecy is measured by the equivocation of the message at the eavesdropper $\frac{1}{n}H(M|W^n)$. The secrecy capacity is defined as the supremum of all rates R where $P_e \leq \epsilon$, and the message is transmitted information-theoretically securely, i.e., $\frac{1}{n}H(M|W^n) \geq \frac{1}{n}H(M) - \epsilon$, in the limit as $\epsilon \to 0$.

We note that encoding and decoding depend only on the input distribution and the conditional marginals of the legitimate and eavesdropper channels. Thus, the secrecy capacity of the channel given in (1)-(2) is equal to the secrecy capacity of the following channel:

$$V_i = A_i U_i + N_{1i} \quad (5)$$

$$W_i = A_i U_i + N_{1i} + \tilde{N}_i \quad (6)$$

where $\tilde{N}_i \sim \mathcal{CN}(0, \sigma_2^2 - \sigma_1^2)$ and $\tilde{N}_i$ is independent of N_{1i}. It is clear that in the channel model of (5)-(6) the eavesdropper's output is a degraded version of the legitimate receiver's output, and $U \to V \to W$. In addition, since $I(U;V) = I(X;Y)$ and $I(U;W) = I(X;Z)$, the secrecy capacity is [2]

$$\begin{aligned} C_s &= \sup_{F \in \mathcal{F}} I(U;V) - I(U;W) \quad (7) \\ &= \sup_{F \in \mathcal{F}} I(X;Y) - I(X;Z) \quad (8) \end{aligned}$$

where F denotes the input distribution drawn from the class of distributions $\mathcal{F}$ which satisfy the given power constraint. Furthermore, the Markov chain $X \to Y \to Z$ holds, because Z is independent of X given V, which follows from the Markov chain $U \to V \to W$, and that the phase of V is independent of X given Y, since the phase of the fading parameter A is uniform and independent of X. As shown by van Dijk [14] for the discrete case, for this continuous case also, we can show that $I(X;Y) - I(X;Z)$ is a concave function of the input distribution, when $X \to Y \to Z$. Thus, to find the secrecy capacity of the channel in (5)-(6), it suffices to solve the convex optimization problem in (8).

Before we determine the secrecy capacity, we note an upper bound on it as:

$$C_s \leq \log\left(1 + \frac{\sigma_h^2 P}{\sigma_1^2}\right) - \log\left(1 + \frac{\sigma_h^2 P}{\sigma_2^2}\right) \quad (9)$$

This upper bound can be derived as follows:

$$\begin{aligned} I(U;V) - I(U;W) = (h(V) - h(W)) \\ - (h(V|U) - h(W|U)) \quad (10) \end{aligned}$$

The first term on the right side of (10) can be upper bounded by using the entropy power inequality:

$$h(V) - h(W) \leq \log\left(\frac{\sigma_h^2 P + \sigma_2^2}{\sigma_h^2 P + \sigma_1^2}\right) \quad (11)$$

and the second term can be lower bounded by noting

$$h(V|U) - h(W|U) \geq h(V|A,U) - h(W|A,U) = \log\frac{\sigma_1^2}{\sigma_2^2} \quad (12)$$

giving the desired upper bound in (9). The inequality in (12) can be derived by noting that $I(V;A|U) \geq I(W;A|U)$. The significance of the upper bound in (9) is that it shows that the secrecy capacity is always finite, even when the power goes to infinity, and also that the secure degrees of freedom of this system is zero as in the cases of non-fading Gaussian wiretap channel and fading Gaussian wiretap channel with perfect CSI.

III. KKT Optimality Conditions

For a channel with continuous alphabet, the supremum in (8) need not be achievable. A sufficient condition for the achievability of the supremum is that there exists a topology on which mutual information is continuous in the input distribution, implying that the difference of two mutual information quantities induced by the same input distribution is also continuous, and the set of allowable input distributions $\mathcal{F}$ is compact. Both of these criteria hold in our case, as was shown in [12, Appendix I]. We solve the maximization in (8) using convex optimization techniques following Smith [11] and Abou-Faycal et al. [12]. The channel input X^* with distribution F^* that achieves the secrecy capacity must satisfy the KKT optimality condition:

$$\begin{aligned} \gamma(x^2 - P) + C_s - \int p_{Y|X}(y|x) \ln\left[\frac{p_{Y|X}(y|x)}{p_Y(y;F^*)}\right] dy \\ + \int p_{Z|X}(z|x) \ln\left[\frac{p_{Z|X}(z|x)}{p_Z(z;F^*)}\right] dz \geq 0, \quad \forall x \in \mathbb{R} \quad (13) \end{aligned}$$

for some $\gamma \geq 0$, which is the Lagrange multiplier due to the average power constraint on the channel input. Furthermore, (13) is satisfied with equality if x lies in the support of X^*.

Note that, in (13), $p_Y(y;F)$ and $p_Z(z;F)$ are the probability distributions of Y and Z, respectively, which are induced by the probability distribution F, of X, i.e.,

$$p_Y(y;F) = \int p_{Y|X}(y|x)\,dF(x) \tag{14}$$

$$p_Z(z;F) = \int p_{Z|X}(z|x)\,dF(x) \tag{15}$$

In the next section, we will examine the implications of the KKT conditions in (13) on the optimum probability distribution for the channel input X.

IV. Characterization of X^*

Theorem 1 *The optimal X^* is discrete with only a finite number of points in any bounded interval.*

Proof: To prove the theorem, we need to rule out the following two cases:

1) The support of X^* contains an interval.
2) X^* is discrete but there exists a bounded interval containing infinitely many points belonging to the support of X^*.

We proceed by contradiction. Therefore, let us assume that either of the two cases 1) or 2) holds. Let E be the support set of X^*. Noting that

$$\int p_{Y|X}(y|x)\ln p_{Y|X}(y|x)\,dy = \ln\left(\frac{1}{\sigma_h^2 x^2+\sigma_1^2}\right) - 1 \tag{16}$$

one can simplify (13) as:

$$f(x) \geq 0, \quad \forall x \in \mathbb{R} \tag{17}$$

with equality if $x \in E$, where $f(x)$ is given by

$$\begin{aligned} f(x) =& \gamma(x^2-P) + C_s + \ln\left(\frac{\sigma_h^2 x^2+\sigma_1^2}{\sigma_h^2 x^2+\sigma_2^2}\right) \\ &+ \int p_{Y|X}(y|x)\ln\left(p_Y(y;F^*)\right)dy \\ &- \int p_{Z|X}(z|x)\ln\left(p_Z(z;F^*)\right)dz \end{aligned} \tag{18}$$

Now, E contains a bounded set S with an infinite number of distinct points. Let S_c be a compact neighbourhood containing S. By the Bolzano-Weierstrass theorem, the set S must have an accumulation point in S_c. We extend $f(x)$ to the complex domain, and by letting $\ln x$ be the principal branch of the logarithm, f is well defined and analytic on the complex plane. The KKT conditions in (17) tell us that, f which is an analytic function on a domain D, is identically zero on a set with an accumulation point in D. The identity theorem tells us that f must be identically zero everywhere on D. More specifically, f must be zero on the entire real line. Thus, the equality in (17) holds, i.e., $f(x)=0$, for all $x \in \mathbb{R}$. Since $X \to Y \to Z$,

$$\begin{aligned} p_{Z|X}(z|x) &= \int p_{Y,Z|X}(y,z|x)dy \qquad (19) \\ &= \int p_{Y|X}(y|x)p_{Z|Y}(z|y)dy \qquad (20) \end{aligned}$$

We use (20) in (18) and exchange the order of integrals using Fubini's theorem, which is permissible since $|\ln p_Z(z;F^*)|$ is bounded by $\alpha+\beta z$ for some constants α and β, as will be shown in (31) and (43). This enables us to rewrite the equation $f(x)=0$, for all $x \in \mathbb{R}$, equivalently as

$$\begin{aligned} \int p_{Y|X}(y|x)g(y)\,dy =& \gamma(P-x^2) - C_s \\ &- \ln\left(\frac{\sigma_h^2 x^2+\sigma_1^2}{\sigma_h^2 x^2+\sigma_2^2}\right), \quad \forall x \in \mathbb{R} \end{aligned} \tag{21}$$

where

$$g(y) = \ln p_Y(y;F^*) - \int p_{Z|Y}(z|y)\ln(p_Z(z;F^*))\,dz \tag{22}$$

Next, we define

$$s = \frac{1}{\sigma_h^2 x^2+\sigma_1^2} \quad \text{and} \quad \Delta = \frac{1}{\sigma_2^2-\sigma_1^2} \tag{23}$$

and get, after some simplification,

$$\begin{aligned} \int e^{-sy}g(y)\,dy =& -\frac{1}{s}\frac{\gamma}{\sigma_h^2}\left(\frac{1}{s}-\sigma_1^2-\sigma_h^2 P\right) - \frac{1}{s}C_s \\ &- \frac{1}{s}\ln\Delta + \frac{1}{s}\ln(s+\Delta) \end{aligned} \tag{24}$$

Now, we recognize the left hand side of (24) as the Laplace transform of $g(y)$, and by taking an inverse Laplace transform of both sides, we get

$$g(y) = -\frac{\gamma}{\sigma_h^2}y - e^{-\Delta y}\ln y - \Delta\int_0^y e^{-\Delta t}\ln t\,dt - K \tag{25}$$

where $K = -\gamma\frac{\sigma_1^2}{\sigma_h^2} - \gamma P + C_s + \ln\Delta + C_E$ is a constant, and C_E is Euler's constant. Thus, we have

$$\begin{aligned} \ln p_Y(y;F^*) =& \int p_{Z|Y}(z|y)\ln p_Z(z;F^*)\,dz - \frac{\gamma}{\sigma_h^2}y \\ &- e^{-\Delta y}\ln y - \Delta\int_0^y e^{-\Delta t}\ln t\,dt - K \end{aligned} \tag{26}$$

Now, we bound each term on the right hand side of (26) to obtain a lower bound on $p_Y(y)$. First, we note

$$\Delta\int_0^y e^{-\Delta t}\ln t\,dt \leq \Delta\int_0^y e^{-\Delta t}\ln y\,dt = (1-e^{-\Delta y})\ln y \tag{27}$$

and thus,

$$e^{-\Delta y}\ln y + \Delta\int_0^y e^{-\Delta t}\ln t\,dt \leq \ln y \tag{28}$$

To bound the first term on the right hand side of (26), we first bound $p_Z(z)$ as,

$$\begin{aligned} p_Z(z) &= \int \frac{1}{\sigma_h^2 x^2+\sigma_2^2}e^{-\frac{z}{\sigma_h^2 x^2+\sigma_2^2}}\,dF(x) \qquad (29) \\ &\geq \int \frac{1}{\sigma_h^2 x^2+\sigma_2^2}e^{-\frac{z}{\sigma_2^2}}\,dF(x) \qquad (30) \\ &\geq \frac{1}{\sigma_h^2 P+\sigma_2^2}e^{-\frac{z}{\sigma_2^2}} \qquad (31) \end{aligned}$$

where we used the fact that $\frac{1}{\sigma_h^2 x^2+\sigma_2^2}$ is convex in x^2, Jensen's inequality and the power constraint. Thus, the first term on the right hand side of (26) can be bounded as:

$$\int p_{Z|Y}(z|y)\ln p_Z(z;F^*)\,dz \geq \ln K_1 - K_2\mathbb{E}[Z|Y=y] \tag{32}$$

where $K_1 = \frac{1}{\sigma_h^2 P+\sigma_2^2}$ and $K_2 = \frac{1}{\sigma_2^2}$.

From (6), $W = V + \tilde{N}$. Denoting the real and imaginary parts of a complex number by subscripts R and I, respectively, we note that,

$$Z = |W|^2 = Y + |\tilde{N}|^2 + 2V_R\tilde{N}_R + 2V_I\tilde{N}_I \tag{33}$$

and therefore,

$$\mathbb{E}[Z|Y=y] = y + (\sigma_2^2 - \sigma_1^2) \tag{34}$$

Using (32), (34) and (28) along with (26), we get,

$$\begin{aligned}\ln p_Y(y;F^*) \geq{}& \ln K_1 - K_2 y - K_2(\sigma_2^2-\sigma_1^2)\\ &- \frac{\gamma}{\sigma_h^2}y - \ln y - K\end{aligned} \tag{35}$$

which implies that

$$p_Y(y) \geq \frac{c_1}{y}e^{-c_2 y}, \quad y \geq 0 \tag{36}$$

for some constants c_1 and c_2. We note that

$$\int_0^1 \frac{c_1}{y}e^{-c_2 y}dy = \infty \tag{37}$$

for any value of c_1 and c_2, and hence $p_Y(y)$ cannot be a valid probability density function and thus we have reached a contradiction. This contradiction implies that the two cases stated at the beginning cannot occur, i.e., the optimum probability distribution cannot contain a continuous interval, or an infinite number of discrete points in a finite interval. Therefore, the optimum probability distribution contains at most a finite number of discrete points in any given finite interval. ■

In the following theorem, we show that, in fact, X^* has a finite number of mass points.

Theorem 2 *The support of X^* has a finite number of points.*

Proof: Again, we proceed by contradiction. Assume that the support of X^* has infinitely many points. Let us denote the mass points by the increasing sequence $\{x_i\}_{i=1}^{\infty}$ and their corresponding probabilities by the sequence $\{p_i\}_{i=1}^{\infty}$. Since, by Theorem 1, there are only finitely many points in any bounded interval, we must have $\lim_{i\to\infty} x_i = \infty$. Then, the output probability is bounded as

$$\begin{aligned}p_Y(y) &= \sum_{i=1}^{\infty} p_i p_{Y|X}(y|x_i) &(38)\\ &\geq p_i p_{Y|X}(y|x_i) &(39)\\ &= \frac{p_i}{\sigma_h^2 x_i^2+\sigma_1^2}e^{-\frac{y}{\sigma_h^2 x_i^2+\sigma_1^2}} &(40)\end{aligned}$$

A similar bound clearly holds for $p_Z(z)$ as well. Also, $p_Y(y)$ can be upper-bounded as,

$$\begin{aligned}p_Y(y) &= \int \frac{1}{\sigma_h^2 x^2+\sigma_1^2}e^{-\frac{y}{\sigma_h^2 x^2+\sigma_1^2}}\,dF(x) &(41)\\ &\leq \int \frac{1}{\sigma_1^2}e^{-\frac{y}{\sigma_h^2 x^2+\sigma_1^2}}\,dF(x) &(42)\\ &\leq \frac{1}{\sigma_1^2}e^{-\frac{y}{\sigma_h^2 P+\sigma_1^2}} &(43)\end{aligned}$$

where we have used the fact that $e^{-\frac{y}{\sigma_h^2 x^2+\sigma_1^2}}$ is concave in x^2, Jensen's inequality and the power constraint.

Now we observe that $f(x)$ in (18) is a continuously differentiable function in x. Also, KKT conditions in (17) imply that $f(x_i) = 0, \forall i \in \mathbb{N}$ and $f(x) \geq 0, \forall x \in \mathbb{R}$. Denoting the derivative of $f(x)$ by $f'(x)$, we must have $f'(x_i) = 0, \forall i$. If not, $f(x)$ will change sign in the neighbourhood of x_i, which is not possible. To compute the derivative of $f(x)$, we note

$$\frac{dp_{Y|X}(y|x)}{dx} = \frac{2\sigma_h^2 x}{(\sigma_h^2 x^2+\sigma_1^2)^2}\left[y - (\sigma_h^2 x^2+\sigma_1^2)\right]p_{Y|X}(y|x) \tag{44}$$

and obtain,

$$\begin{aligned}f'(x) ={}& 2\gamma x + \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_1^2} - \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_2^2}\\ &+ \frac{2\sigma_h^2 x}{(\sigma_h^2 x^2+\sigma_1^2)^2}\int y p_{Y|X}(y|x)\ln\left(p_Y(y)\right)dy\\ &- \frac{2\sigma_h^2 x}{(\sigma_h^2 x^2+\sigma_1^2)}\int p_{Y|X}(y|x)\ln\left(p_Y(y)\right)dy\\ &- \frac{2\sigma_h^2 x}{(\sigma_h^2 x^2+\sigma_2^2)^2}\int z p_{Z|X}(z|x)\ln\left(p_Z(z)\right)dz\\ &+ \frac{2\sigma_h^2 x}{(\sigma_h^2 x^2+\sigma_2^2)}\int p_{Z|X}(z|x)\ln\left(p_Z(z)\right)dz\end{aligned} \tag{45}$$

Using the bounds in (40) and (43) to bound the different terms in (45), we obtain

$$\begin{aligned}f'(x) \geq{}& 2\gamma x + \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_1^2} - \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_2^2} - \frac{2\sigma_h^2 x}{\sigma_h^2 x_i^2+\sigma_2^2}\\ &+ \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_1^2}\ln\left(\frac{p_i}{\sigma_h^2 x_i^2+\sigma_1^2}\right) - \frac{4\sigma_h^2 x}{\sigma_h^2 x_i^2+\sigma_1^2}\\ &- \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_1^2}\ln\frac{1}{\sigma_1^2} + \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_2^2}\ln\left(\frac{1}{\sigma_h^2 x_i^2+\sigma_2^2}\right)\\ &- \frac{2\sigma_h^2 x}{\sigma_h^2 x^2+\sigma_2^2}\ln\frac{1}{\sigma_2^2} + \frac{2\sigma_h^2 x}{\sigma_h^2 P+\sigma_1^2} + \frac{4\sigma_h^2 x}{\sigma_h^2 P+\sigma_2^2}\end{aligned} \tag{46}$$

Therefore, we have

$$f'(x_i) \geq \left(2\gamma + \frac{2\sigma_h^2}{\sigma_h^2 P+\sigma_1^2} + \frac{4\sigma_h^2}{\sigma_h^2 P+\sigma_2^2}\right)x_i + o(x_i) \tag{47}$$

where $o(x)$ denotes a function such that $o(x) \to 0$ as $x \to \infty$. By our assumption, $x_i \to \infty$ as $i \to \infty$. Thus, (47) implies that $f'(x_i) \to \infty$ as $i \to \infty$ which is a contradiction, since, $f'(x_i) = 0$, for every i. We conclude, therefore, that the

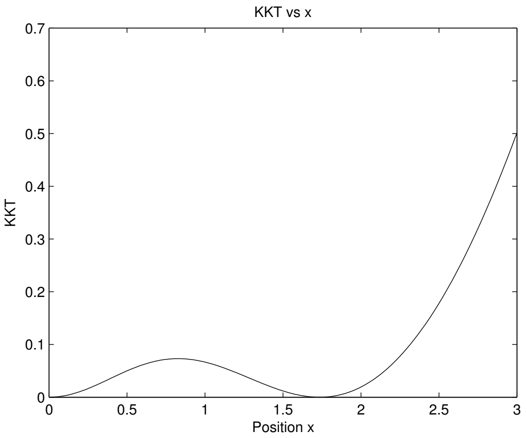

Fig. 1. An optimal distribution satisfying the KKT conditions with $P = 0.1$, $\sigma_h = \sigma_1 = 1$, $\sigma_2 = 2$, $\gamma = 0.2461$, $C_s = 0.03$ and $F(x) = 0.9668\delta(x) + 0.0332\delta(x - 1.7348)$.

support of the optimal input distribution has a finite number of points. ■

V. Numerical Results

In this section, we present simple numerical examples to verify and illustrate the results of this paper. Fig. 1 shows an example of how the KKT conditions are satisfied for a particular value of power P. The plot shows that there are two mass points, one at 0 and the other at 1.7348, with probabilities 0.9668 and 0.0332, respectively. The secrecy capacity for this case is 0.03 bits per channel use.

Fig. 2 shows how the positions of the optimum probability mass points change with power. Note that there is always a mass point at zero. As the power increases, the optimum probability distribution has more and more mass points. At the transitions, where a new mass point is introduced, the numerical algorithm becomes unstable, nevertheless, it seems that the mass points originate far from the origin with very low probabilities (as seen in Fig. 3), then come closer towards the origin before receding away again with increasing power. Fig. 3 shows the probabilities of the corresponding mass points. As expected, at very low power, the probability of the point at zero is high, and it decreases as power is increased. The probabilities stabilize asymptotically.

VI. Conclusion

We considered the fast Rayleigh fading wiretap channel with coherence time of one symbol duration. We proved that the optimal input distribution that achieves the secrecy capacity is discrete with finite number of mass points.

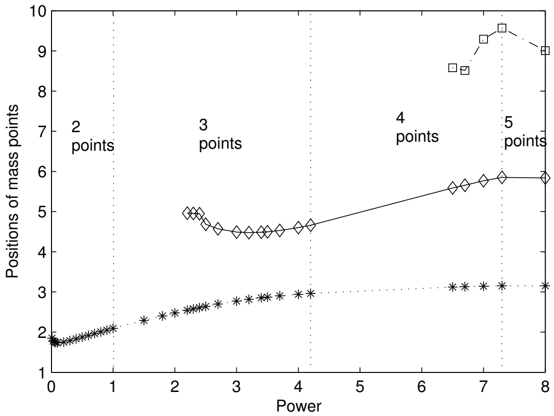

Fig. 2. The position of the mass points versus power.

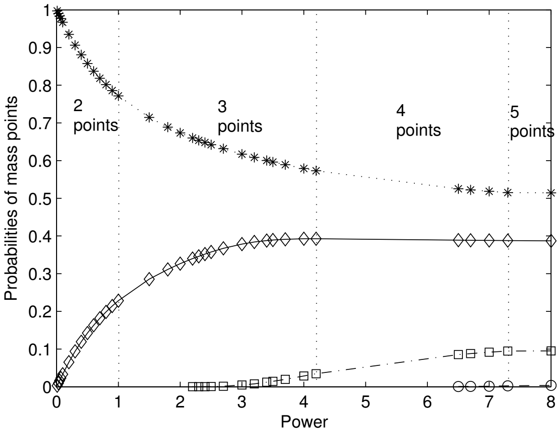

Fig. 3. The probabilities of the mass points versus power.

References

[1] C. E. Shannon. Communication theory of secrecy systems. *The Bell System Technical Journal*, 28(4):656–715, October 1949.

[2] A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, 54(8):1355–1387, October 1975.

[3] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978.

[4] S. Leung-Yan-Cheong and M. Hellman. The Gaussian wire-tap channel. *IEEE Transactions on Information Theory*, 24(4):451–456, July 1978.

[5] Y. Liang and H. V. Poor. Secure communication over fading channels. In *the 44th annual Allerton Conference on Communication, Control, and Computing*, September 2006.

[6] Z. Li, R. D. Yates, and W. Trappe. Secrecy capacity of independent parallel channels. In *the 44th annual Allerton Conference on Communication, Control, and Computing*, September 2006.

[7] Y. Liang, H. V. Poor, and S. Shamai. Secure communication over fading channels. *IEEE Transactions on Information Theory*, 54(6):2470–2492, June 2008.

[8] Z. Li, R. D. Yates, and W. Trappe. Secrecy capacity of independent parallel channels. In R. Liu and W. Trappe, editors, *Securing Wireless Communications at the Physical Layer*, pages 1–18. Springer US, 2010.

[9] P. K. Gopala, L. Lai, and H. El Gamal. On the secrecy capacity of fading channels. *IEEE Transactions on Information Theory*, 54(10):4687–4698, October 2008.

[10] Z. Li, R. D. Yates, and W. Trappe. Achieving secret communication for fast Rayleigh fading channels. *IEEE Transactions on Wireless Communications*, 9(9):2792–2799, September 2010.

[11] J. G. Smith. The information capacity of amplitude and variance-constrained scalar Gaussian channels. *Information and Control*, 18(3):203–219, April 1971.

[12] I. C. Abou-Faycal, M. D. Trott, and S. Shamai. The capacity of discrete-time memoryless Rayleigh-fading channels. *IEEE Transactions on Information Theory*, 47(4):1290–1301, May 2001.

[13] A. Agrawal, Z. Rezki, A. J. Khisti, and M. Alouini. Noncoherent capacity of secret-key agreement with public discussion. *IEEE Transactions on Information Forensics and Security*, 6(3):565–574, September 2011.

[14] M. van Dijk. On a special class of broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 43(2):712–714, March 1997.