Non-Coherent Capacity of Secret-Key Agreement with Public Discussion

Anurag Agrawal, Zouheir Rezki,Ashish Khisti and Mohamed-Slim Alouini,

***Abstract*—We study the Rayleigh fading non-coherent capacity of secret-key agreement with public discussion, where neither the sender nor the receivers have access to instantaneous channel state information (CSI) of any channel. We present two results. At high Signal-to-Noise Ratio (SNR), the secret-key capacity is bounded in SNR, regardless of the number of antennas at each terminal. Second, for a system with a single antenna at both the legitimate and the eavesdropper terminals and an arbitrary number of transmit antennas, the secret-key capacity-achieving input distribution is discrete, with a finite number of mass points. Numerically we observe that at low-SNR, the capacity achieving distribution has two mass points with one of them at the origin.**

***Index Terms*—Non-coherent capacity, secret-key agreement, Rayleigh fading channels, information theoretic security, Karush-Kuhn-Tucker (KKT) condition, discrete input distribution.**

I. Introduction

Information theoretic secret-key agreement provides provably secure mechanisms for generating secret-keys between two or more legitimate terminals. In such protocols, the legitimate terminals need to have access to a source of correlated randomness e.g., communication channels or correlated sources [1], [2]. Furthermore a discussion channel of unlimited capacity is also available for communication, but is public to the wiretapper. The legitimate terminals distill a common secret-key that satisfies an equivocation constraint with respect to the eavesdropper.

The present paper studies capacity limits of secret-key agreement when the underlying channel from the sender to the receiver and the eavesdropper are modeled as independent identically distributed (i.i.d.) Rayleigh fading. We further assume the non-coherent model i.e., the instantaneous channel state information is not known to either of the terminals. The channel statistics are however globally known.

Note that for our proposed channel model the outputs at the legitimate receiver and the eavesdropper are conditionally independent given the channel input. A class of discrete memoryless channel models with this property was studied in [1], [2] and a single-letter capacity expression was characterized. In particular a *source-emulation* strategy was shown to be optimal — the sender generates a discrete memoryless source, then transmits it over the channel to generate correlated sources at the two terminals and then the legitimate terminals distill a common key as in the source model. While their result can be extended using standard techniques to the (continuous-valued) Rayleigh fading channels studied in this work, finding the optimizing distribution is difficult in general. In the present work we show the following two properties (1) Unlike the case without secrecy constraint where the capacity scales as $\log\log(P)$ at high-SNR, the secret-key capacity is bounded in SNR, regardless of the number of antennas at each terminal and (2) The capacity achieving distribution is discrete with a finite number of mass points, for the MISO case.

In related works, [3] studies the secret-key agreement over Rayleigh fading channels for the case of receiver CSI and establishes that a Gaussian input distribution maximizes the secret-key capacity. Other works, see e.g.[4], study the problem of generating shared secret keys using channel reciprocity instead of public discussion. This approach is not considered in the present paper.

Anurag Agrawal (anuragagrawal2006@gmail.com) is with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India. Zouheir Rezki and Mohamed-Slim Alouini are with King Abdullah University of Science and Technology (KAUST) {zouheir.rezki;slim.alouini}@kaust.edu.sa, Thuwal, KSA. Ashish Khisti (akhisti@comm.utoronto.ca) is with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. This work was conducted when Anurag Agrawal and Zouheir Rezki were visiting University of Toronto

II. The Channel Model

Consider a Discrete-Time Memoryless Channel (DMC) consisting of a transmitter, a legitimate receiver and an eavesdropper, with n_T,n_R and n_E antennas, respectively. The outputs at both the legitimate destination and the eavesdropper, at time $i = 1, \ldots, L$, are expressed, respectively by:

$$\begin{cases} \mathbf{y}(i) = \mathbf{H}(i)\mathbf{x}(i) + \mathbf{v}(i) \\ \mathbf{z}(i) = \mathbf{G}(i)\mathbf{x}(i) + \mathbf{w}(i) \end{cases} \tag{1}$$

where $\mathbf{x}(i) \in \mathbb{C}^{n_t}$ is the transmitted signal, and $\mathbf{H}(i) \in \mathbb{C}^{n_R \times n_T}$, $\mathbf{G}(i) \in \mathbb{C}^{n_E \times n_T}$ represent the main channel and the eavesdropper channel gains, respectively; and $\mathbf{v}(i) \in \mathbb{C}^{n_R}$, $\mathbf{w}(i) \in \mathbb{C}^{n_E}$ are circularly symmetric white Gaussian noises with covariance matrices $\mathbb{E}[\mathbf{v}(i)\mathbf{v}(i)^\dagger] = \mathrm{I}_{n_R}$ and $\mathbb{E}[\mathbf{w}(i)\mathbf{w}(i)^\dagger] = \mathrm{I}_{n_E}$. We assume that $\mathbf{H}(i)$ and $\mathbf{G}(i)$ have independent and identically-distributed Gaussian entries with zero-mean and unit-variance. We assume that the Channel State Information (CSI) is not available at any terminal. That is, the transmitter, the legitimate receiver and the eavesdropper do not have access to the instantaneous channel realizations $\mathrm{H}(i)$ and $\mathrm{G}(i)$; but are aware of their statistic. The source is constrained according to a short-term average power constraint:

$$\mathbb{E}[\mathbf{x}(i)^\dagger \mathbf{x}(i)] = \mathrm{tr}(\mathrm{Q}_x) \leq P \tag{2}$$

for all $i = 1, \ldots, L$. Since the channel defined in (1) is i.i.d. we may drop the time index i in the sequel for convenience.

III. Secret-Key Capacity

In [1, Theorem 2], a single-letter formula of key-capacity has been established and is given by:

$$C = \sup_{F(\mathbf{x}) \in \mathcal{F}} I(\mathbf{x}; \mathbf{y}|\mathbf{z}), \tag{3}$$

where $\mathcal{F}$ is the set of all possible distribution functions $F(\mathbf{x})$ that satisfy the average power constraint. Note that in our model of interest $\mathbf{y}$ and $\mathbf{z}$ are conditionally independent given $\mathbf{x}$ and the distribution is given by [5]:

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})=\frac{1}{\pi^{n_R}(1+\|\mathbf{x}\|^2)^{n_R}}\ \exp\left(\frac{-\|\mathbf{y}\|^2}{1+\|\mathbf{x}\|^2}\right) \quad (4)$$

$$p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x})=\frac{1}{\pi^{n_E}(1+\|\mathbf{x}\|^2)^{n_E}}\ \exp\left(\frac{-\|\mathbf{z}\|^2}{1+\|\mathbf{x}\|^2}\right), \quad (5)$$

where $\|\cdot\|$ denotes the Euclidean norm of a vector. Noting that $p_{\mathbf{y}|\mathbf{x}}$ and $p_{\mathbf{z}|\mathbf{x}}$ depend on $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{x}$ only through their norms, then by letting $X=\|\mathbf{x}\|$, $Y=\|\mathbf{y}\|^2$ and $Z=\|\mathbf{Z}\|^2$, $I(\mathbf{x};\mathbf{y}|\mathbf{z})$ can be formulated as follows:

$$\begin{aligned} I(\mathbf{x};\mathbf{y}|\mathbf{z})&=I(\mathbf{x};\mathbf{y},\mathbf{z})-I(\mathbf{x};\mathbf{z})\\ &=I(X;Y,Z)-I(X;Z)\\ &=I(X;Y|Z) \end{aligned} \quad (6)$$

where (6) follows from the fact that Y and Z are sufficient statistics that preserve the mutual information and because $p_{\mathbf{y}|\mathbf{x}}$ and $p_{\mathbf{z}|\mathbf{x}}$ depend only on X. This result is summarized in the following lemma.

Lemma 1: The non-coherent secret-key capacity of the channel model (1) described above is given by:

$$C=\sup_{F(x)\in\mathcal{F}} I(X;Y|Z). \quad (7)$$

Lemma 1 states that secret-key communication is conveyed over the norms X, Y and Z. This property will be used in the sequel. As verified in Appendix A, the capacity can be expressed as

$$C=\sup_{F(x)\in\mathcal{F}}\iiint p_{Y|X}(y|x)p_{Z|X}(z|x)\ln\left[\frac{p_{Y|X}(y|x)}{p_{Y|Z}(y|z;F)}\right]dydzdF(x) \quad (8)$$

provided conditional densities $p_{Y|X}(y|x)$ and $p_{Z|X}(y|x)$ exist. The later conditions is however guaranteed by the nature of our channel model (1) and in particular from (4) and (5),

$$\begin{aligned} p_{Y|X}(y|x)&=\frac{y^{(n_R-1)}}{\Gamma(n_R)\,(1+x^2)^{n_R}}\ \exp\left(\frac{-y}{1+x^2}\right),\\ p_{Z|X}(z|x)&=\frac{z^{(n_E-1)}}{\Gamma(n_E)\,(1+x^2)^{n_E}}\ \exp\left(\frac{-z}{1+x^2}\right). \end{aligned} \quad (9)$$

IV. Capacity Results At A High-SNR Regime

In this section, we analyze the non-coherent secret key capacity asymptotically at high-SNR. Our result is rather negative as it establishes the non-efficiency of communication over this channel at high-SNR. Theorem 1 below formalizes this result.

Theorem 1: At high-SNR, the non-coherent secret-key capacity of the channel model (1) described above is given by:

$$C(P)=O(1) \quad (10)$$

Proof: By Lemma 1, it can be seen that the channel (1) is equivalent, from a capacity perspective, to the multiplicative channel:

$$\begin{cases} Y=(1+X^2)W_1\\ Z=(1+X^2)W_2 \end{cases} \quad (11)$$

where it follows via (9) that W_1 and W_2 are Gamma-distributed random variables, mutually independent and also independent of X, and with probability density functions (p.d.f.):

$$p_{W_1}(w_1)=\frac{w_1^{n_R-1}}{\Gamma(n_R)}\ \exp(-w_1) \quad (12)$$

$$p_{W_2}(w_2)=\frac{w_2^{n_E-1}}{\Gamma(n_E)}\ \exp(-w_2). \quad (13)$$

Now, by letting $Y_1=\ln(Y)$, $Z_1=\ln(Z)$, $D_1=\ln(1+X^2)$, $N_1=\ln(W_1)$ and $N_2=\ln(W_2)$, and by applying the log function on both sides of (11), the following channel is obtained:

$$\begin{cases} Y_1=D_1+N_1\\ Z_1=D_1+N_2 \end{cases} \quad (14)$$

Since the log function is a one-to-one transformation that does not entail any capacity loss, the secret-key capacity can be bounded as follows:

$$\begin{aligned} I(X;Y|Z)&=I(D_1;Y_1|Z_1)\\ &=h(Y_1|Z_1)-h(Y_1|Z_1,D_1)\\ &=h(Y_1-Z_1|Z_1)-h(Y_1|Z_1,D_1)\\ &\le h(Y_1-Z_1)-h(N_1) \quad (15)\\ &=h(N_1-N_2)-h(N_1). \quad (16) \end{aligned}$$

Hence, it remains to show that the right hand side (RHS) of (16) is bounded. In particular it suffices to show that $E[N_1^2]<\infty$ and $E[N_2^2]<\infty$, since a Gaussian distribution upper bounds the differential entropy of any continuous distribution.

$$\begin{aligned} \mathbb{E}\left[N_1^2\right]&=\mathbb{E}\left[\left(\ln(W_1)\right)^2\right]\\ &\le\mathbb{E}\left[\left(\ln(W_1)\right)^2|W_1\le 1\right]\Pr(W_1\le 1)\\ &\quad+\mathbb{E}\left[\left(\ln(W_1)\right)^2|W_1>1\right]\Pr(W_1>1)\\ &\le\mathbb{E}\left[\frac{1}{W_1}|W_1\le 1\right]\Pr(W_1\le 1)\\ &\quad+\mathbb{E}[W_1|W_1>1]\Pr(W_1>1) \quad (17)\\ &\le\mathbb{E}\left[\frac{1}{W_1}\right]+\mathbb{E}[W_1]\\ &=\frac{1}{n_R-1}+n_R \quad (18) \end{aligned}$$

In the above derivation, (17) follows from the fact that $x\left(ln(x)\right)^2\le 1$ for any $x\in[0,1]$, whereas (18) holds when $n_R>1$ because $\frac{1}{W_1}$ follows inverse Gamma distribution with parameters $\beta=1$ and $\alpha=n_R$. For $n_R=1$ we can see that the random variable N_1, being the log of an exponential random variable, has density function $p_{N_1}(x)=e^x e^{-e^x}$ with finite mean and variance. By a similar argument, $\mathbb{E}\left[N_2^2\right]$ is finite, and thus $Var(N_1-N_2)=\mathbb{E}\left[N_1^2\right]+\mathbb{E}\left[N_2^2\right]$ is also finite. Now, we have an upper bound on $I(X;Y|Z)$ that is itself bounded irrespective to the power P. We conclude that the secret-key capacity is asymptotically bounded at high-SNR. ■

V. The Karush-Kuhn-Tucker (KKT) condition

In this section, following [6], a necessary and sufficient condition for optimality in secret-key agreement settings is established. Our proof relays on two steps. First, we show that the supremum in (8) is achieved (Lemma 2). Next, we argue that $I(X;Y|Z)$ is also weak differentiable in F over $\mathcal{F}$ (Lemma 3). The KKT follows then by the concavity of a modified objective function that includes the power constraint. Although we focus in our proof on the MISOSE case, i.e., $n_R = n_E = 1$, we believe that our framework can be extended in a straightforward manner to encompass a channel where n_R and n_E are arbitrary.

Lemma 2: The supremum in (8) is achievable by at least one F, say F^*, belonging to $\mathcal{F}^+$, where $\mathcal{F}^+$ is the set of all nonnegative input distributions that meet the power constraint.

Proof: A sufficient condition for the supremum in (8) to exist, is that the mutual information $I(X;Y|Z)$ be weak continuous in F and the set $\mathcal{F}^+$ be weak compact. That $\mathcal{F}^+$ is weak compact follows from [6, Appendix 1.A], whereas the proof of weak continuity of $I(X;Y|Z)$ in F is reported in the full version of this paper [7]. ∎

Lemma 3: $I(X;Y|Z)$ is weak differentiable and concave in F over $\mathcal{F}^+$.

Proof: The proof of weak differentiability is presented in [7]. The concavity of $I(X;Y|Z)$ in F, follows from [8, Fact 2]. ∎

Now, from Lemma 2 and Lemma 3, a necessary and sufficient condition for optimality, referred to as the KKT condition, can be obtained as stated below.

Theorem 2: For the channel given by (1) described above, where $n_R = n_E = 1$, an input random variable X^* with distribution function F^* achieves the secret-key capacity C if and only if there exists a $\gamma \geq 0$ such that,

$$\gamma(x^2 - P) + C - \iint p_{Y|X}(y|x)p_{Z|X}(z|x)\ln\left[\frac{p_{Y|X}(y|x)}{p_{Y|Z}(y|z;F^*)}\right]dy\,dz \geq 0 \tag{19}$$

for all x, with equality if x belongs to the support of X^*.

Proof: The proof is presented in Appendix [7]. ∎

Using (12), (13) and letting $s = \frac{1}{1+x^2}$ with $s \in [0,1]$, (19) can be expressed as:

$$\gamma\left(\frac{1}{s} - 1 - P\right) + C - \ln(s) + 1 + \iint s^2 e^{-s(y+z)}\ln\left[p_{Y|Z}(y|z;F^*)\right]dy\,dz \geq 0. \tag{20}$$

VI. Characterization of X^*

Here we follow [6] and [9] to use the Kuhn-Tucker condition (19) to prove that X^* is discrete. Although our framework parallels these previous works, several modifications are necessary to account for the conditional mutual information. The existence of a secret-key capacity achieving input implies that X^* should satisfy one of the following properties:

1) Its support contains an interval;
2) It is discrete, with an infinite number of mass points in some bounded interval;
3) It is discrete and infinite, but with only a finite number of mass points on any bounded interval;
4) It is discrete with a finite number of mass points.

Now, let us assume that 1) or 2) holds and define the function $H(z)$ by:

$$H(z_c) = \gamma\left(\frac{1}{z_c} - 1 - P\right) + C - \ln(z_c) + 1 + \iint z_c^2 e^{-z_c(y+z)}\ln\left[p_{Y|Z}(y|z;F^*)\right]dydz \tag{21}$$

for all z_c belonging to the set of complex numbers, where $\ln(\cdot)$ is the principal branch of the logarithm. We note that $H(z_c)$ is analytic over the domain D defined by $\Re(z_c) > 0$. We now make the following observations:

- From our assumption it is evident that there exists an A such that the support of X^* contains infinitely many points in $[0, A]$ or equivalently the support of S^* contains an infinite set of distinct points $S_s \subseteq [1/(1+A^2), 1]$,
- The interval $[1/(1+A^2), 1]$ is compact, hence by Bolzano-Weierstrass theorem S_s has an accumulation point in $[1/(1+A^2), 1] \subset [0,1]$,
- From the Kuhn-Tucker condition (20), $H(z_c) = 0$ on the support of S^* and thus on S_s.

Hence we have an analytic function over D that vanishes on a set having an accumulation point of in D. From the identity theorem [10], we conclude that $H(z_c) = 0$ over the whole D and in particular, over $z_c \in [0,\infty)$. Consequently, (21) can be written as:

$$\iint e^{-z_c(y+z)}\ln\left[p_{Y|Z}(y|z;F^*)\right]dydz = \frac{-1}{z_c^2}\left[\gamma(\frac{1}{z_c} - 1 - P) + C - \ln(z_c) + 1\right], \tag{22}$$

where $z_c \in D$. We next show that there cannot be a valid conditional probability density function $p_{Y|Z}(\cdot|\cdot)$ that satisfies (22). This is done by multiplying both sides of (22) by z_c^2 and by taking the limit $z_c \to \infty$ we show that while the right hand side diverges, the left hand side remains bounded. Towards this end we show the following.

Lemma 4: For each z_c with $\Re(z_c) > 0$ we have that

$$\iint e^{-z_c(y+z)}\ln\left[p_{Y|Z}(y|z;F^*)\right]dydz = I_1(z_c) - I_2(z_c), \tag{23}$$

where $I_1(\cdot)$ and $I_2(\cdot)$ denote the Laplace transforms of the following

$$I_1(\cdot) = \mathcal{L}(w\ln f(w)) \tag{24}$$

$$I_2(\cdot) = \mathcal{L}\left(\int_0^w \ln g(\tau)d\tau\right) \tag{25}$$

where we define

$$f(t) = \int_0^1 s^2 e^{-st} dF^*(s) \tag{26}$$

$$g(t) = \int_0^1 s e^{-st} dF^*(s). \tag{27}$$

Furthermore $I_1(z_c)$ and $I_2(z_c)$ are well defined for all z_c with $\Re(z_c) > 0$.

Proof: The proof of Lemma 4 is provided in Appendix B. ∎

Note that $f(t)$ and $g(t)$ in (24) and (25) are infinitely differentiable decreasing functions over the positive real axis. Furthermore,

$$\lim_{t\to 0} g(t) = g(0) = \mathbb{E}\left[\frac{1}{(1+X^2)}\right] \quad \in \quad [0,1] \tag{28}$$

Similarly, we have:

$$\lim_{t\to 0} f(t) = \mathbb{E}\left[\frac{1}{(1+X^2)^2}\right] \quad \in \quad [0,1] \tag{29}$$

$$\lim_{t\to 0} tf'(t) = 0. \tag{30}$$

Finally, multiplying both sides of (22) by z_c^2 and taking the limit as $z_c \to \infty$, we show that the RHS goes to infinity whereas LHS is finite. To see this, we first recall that:

$$\lim_{z_c\to\infty} z_c^2 LHS = \lim_{z_c\to\infty} \left[z_c^2\left(I_1(z_c) - I_2(z_c)\right)\right]. \tag{31}$$

Then, the limit of the second term on the RHS of (31) is equal to:

$$\begin{aligned}\lim_{z_c\to\infty} z_c^2 I_2(z_c) &= \lim_{z_c\to\infty}\left[z_c\int_0^\infty e^{-z_c w}\ln[g(w)]dw\right] \\ &= \lim_{w\to 0}[\ln(g(w))] \qquad (32)\\ &= \ln\left(\mathbb{E}\left[1/\left(1+X^2\right)\right]\right), \qquad (33)\end{aligned}$$

where (32) follows by the Initial Value Theorem [11] and (33) is obtained from (28). Next, we note that the second derivative $(w\ln(f(w)))''$ exists for all positive w, that $\lim_{w\to 0}[w\ln(f(w))] = 0$ and that $\lim_{w\to 0}\left[(w\ln(f(w)))'\right] = \ln\left(\mathbb{E}\left[1/\left(1+X^2\right)^2\right]\right)$. Hence, applying the identity,

$$\int_0^\infty e^{-z_c t}p''(t)dt = z_c^2\int_0^\infty e^{-z_c t}p(t)dt - z_c p(0) - p'(0) \tag{34}$$

to the function $w\ln(f(w))$ and taking the limit as $z_c \to \infty$ on both sides of (34) yield,

$$0 = \lim_{z_c\to\infty}\left[z_c^2\left(I_1(z_c)\right)\right] - \ln\left(\mathbb{E}\left[1/\left(1+X^2\right)^2\right]\right), \tag{35}$$

which confirms that $\lim_{z_c\to\infty}\left[z_c^2\left(I_1(z_c)\right)\right]$ is finite too. Therefore, the limit in (31) exists and is finite. This implies that (22) does not hold for all $z_c \in D$. But, this contradicts our initial assumption that either 1) or 2) holds. Consequently, neither 1) nor 2) can happen. We are then left with 3) and 4) as the only possibilities.

Let us assume that 3) holds. In this case we argue that the Lagrange multiplier in the KKT condition (20) is zero and in turn obtain a contradiction. Since X^* has infinitely many mass points and only finitely many in any bounded interval, S^* has an accumulation point only at zero and its support can thus be written as a sequence $\{s_i\}$ converging to zero. Let $\Pr\{S^* = s_i\} = p_i$. Then we have:

$$\begin{aligned}p_{Y|Z}(y|z;F^*) &= \frac{\sum_{i=1}^\infty p_i p_{Y|S^*}(y|s_i)p_{Z|S^*}(z|s_i)}{\sum_{i=1}^\infty p_i p_{Z|S^*}(z|s_i)} \qquad (36)\\ &= \frac{\sum_{i=1}^\infty p_i s_i^2 e^{-s_i(y+z)}}{\sum_{i=1}^\infty p_i s_i e^{-s_i z}}. \qquad (37)\end{aligned}$$

Note that the denominator of (37) is smaller $E[S]$ which is less than 1 (c.f. (28)). This implies that $p_{Y|Z}(y|z;F^*) > p_i s_i^2 e^{-s_i(y+z)}$ for all $y, z \geq 0$ and all $i = 1, 2, \ldots$. Consequently, we have:

$$\begin{aligned}&\iint s^2 e^{-s(y+z)}\ln\left[p_{Y|Z}(y|z;F^*)\right]dydz \\ &> \iint s^2 e^{-s(y+z)}\ln\left[p_i s_i^2 e^{-s_i(y+z)}\right]dydz \qquad (38)\\ &= \ln\left(p_i s_i^2\right) - \frac{2s_i}{s} \qquad (39)\end{aligned}$$

Now, using (39), bound the LHS of (20) as follows:

$$\begin{aligned}\text{LHS} &> \gamma\left(\frac{1}{s} - 1 - P\right) + C - \ln(s) + 1 + \ln(p_i s_i^2) - \frac{2s_i}{s} \\ &= \frac{\gamma - 2s_i}{s} + o\left(\frac{1}{s}\right) \qquad (40)\end{aligned}$$

Where the $o(\cdot)$ term applies when $s \to 0$ for a fixed i. Now, if $\gamma > 2s_i$ then (40) goes to infinity as $s \to 0$, but the LHS of (20) should be zero on the support of S^* which by our initial assumption, contains a point of accumulation at 0. Hence, $\gamma \leq 2s_i$. As this is true for all s_i, and $s_i \to 0$, we see that $\gamma \leq 0$. As the Lagrange multiplier is non-negative, we conclude that $\gamma = 0$. Then from (20) we get,

$$C - \ln(s_i) + 1 + \iint s_i^2 e^{-s_i(y+z)}\ln\left[p(y|z;F^*)\right]dydz = 0 \tag{41}$$

for all s_i. Now taking the limit at $s_i \to 0$ on both sides of (41), we see that the integrand on the LHS tends toward 0, which implies that

$$C = \lim_{s_i\to 0}\ln(s_i) - 1 \tag{42}$$

and consequently the capacity C goes to $-\infty$, which contradicts $C \geq 0$. Hence, the assumption 3) is ruled out as well. Therefore, the optimum input distribution must be discrete with a finite number of mass points which we wanted to prove.

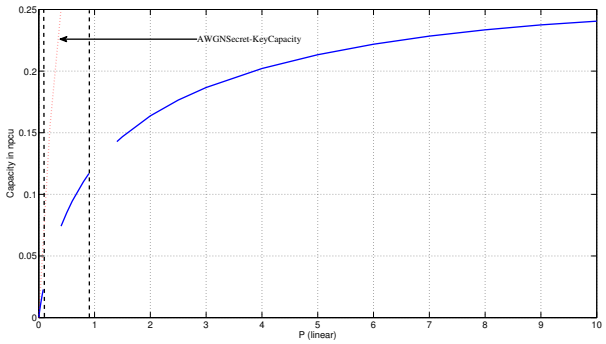

Fig. 1. The SISOSE secret-key capacity versus the SNR value P.

VII. Numerical Results

We employed the Gauss-Laguerre quadrature method to evaluate all the concerned integrals in obtaining capacity-achieving input distributions. We obtain some useful insights related to the variation of the number of mass points and their respective probabilities with the SNR. Furthermore we exploit the variation of KKT, a necessary and sufficient condition for optimality, with x to exactly predict the location of a new mass point and to validate the optimal input distribution for a given SNR value. It has been observed that there exists a mass point

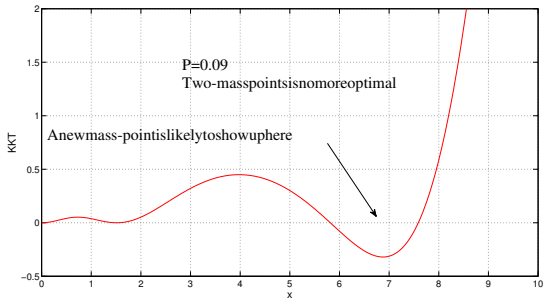

Fig. 2. The value of the LHS of the KKT condition (19) versus x at SNR=0.09.

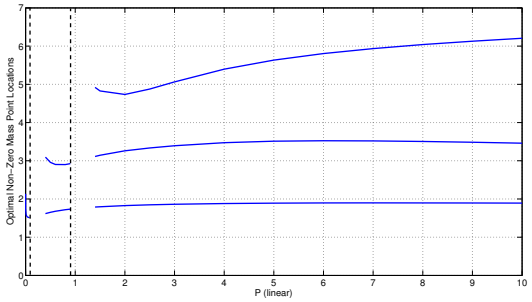

Fig. 3. The optimal non zero mass point locations versus the SNR value P.

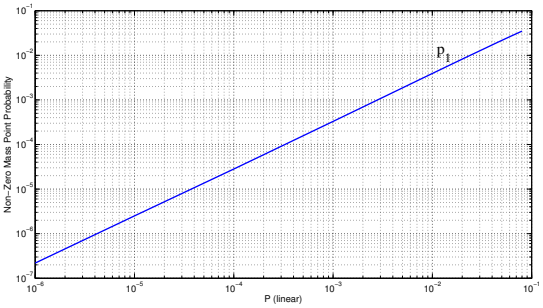

Fig. 4. The non zero mass point probability versus the SNR value P at very low SNR region.

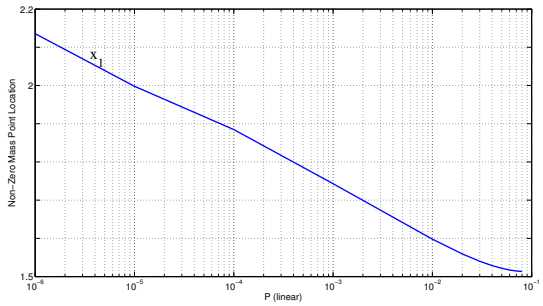

Fig. 5. The non zero mass point location versus the SNR value P at very low SNR region .

at the origin for all SNR values. Figure 1 represents the non-coherent secret-key capacity in nats per channel use (npcu) in function of SNR (average power constraint in Joules per second). As shown in Fig. 1, the secret-key capacity shown in solid blue is monotonically increasing in SNR. In this figure, discontinuities of the capacity plot can be observed. Indeed, theses regions represent zones where a new mass point is about to appear and where numerical optimization becomes very instable to an extent that the results obtained in these zones do not fulfill the KKT condition. Furthermore, it may be seen in Fig. 1, that the capacity is bounded at high-SNR in full agreement with Theorem 1. Also shown in Fig.1 as a benchmark, is the Additive White Gaussian Noise (AWGN) channel secret-key capacity (in dashed red). While the gap between the two plots is marginal at very-low SNR (below 0.5 J/s), the AWGN secret-key capacity prevails remarkably as SNR increases.

In Fig. 2, the LHS of the KKT condition (19) is plotted versus x for an SNR$= 0.09$ and where numerical optimization was set to two mass point input distributions ($N = 2$). From Fig.2, it can be seen that although the KKT is null in two points, the results obtained is not optimal since the plot goes below zero for a certain value of x, suggesting that a new mass point is more likely to appear. In order to confirm our claim, we set $N = 3$ in our optimization problem and increase the power constraint away from this instable zone, to find that three mass point is in fact optimal and that a new mass point shows up around $x = 3$ at approximatively $P = 0.4$. This explains the discontinuities in our capacity plot as depicted by Fig. 1. We conjuncture that a new mass point shows up first at $x \to \infty$ and then decreases as SNR increases. Note that this peculiar behavior of the optimal input distribution has also been observed previously in non-coherent fading channels without secrecy [6].

Figure 3 depicts optimal non zero mass point locations versus SNR. Likewise in Rayleigh fading channel without secrecy, two mass point is optimal at low SNR (below $P = 0.08$). As SNR increases, the number of mass point increases gradually to be $N = 3$ for P below 0.9 and then $N = 4$ for P below 10. At high-SNR, we observe that the optimal distribution has 3 non zero mass points and more interestingly as SNR increases, only the biggest of the three tends to increase whereas the two others of lower values tend to attain constant values of approximately 1.9 and 3.5, respectively.

Finally, the non zero mass point probability versus P is shown in Fig. 4 at low-SNR, where it can be seen that the non zero mass point probability seems to increase almost linearly with SNR. On the other hand, the non zero mass point location versus P is also displayed in Fig. 5 at low-SNR, where it can be observed that as SNR increases the non zero mass point decreases in magnitude.

VIII. Conclusion

The secret-key capacity under an average power constraint of a Rayleigh fading channel, where the instantaneous CSI is not available at any terminal, has been studied. When the legitimate receiver and the eavesdropper have each one antenna, i.e., the MISOSE ($n_R = n_E = 1$) setting, it has been shown that the capacity-achieving input distribution is discrete with a finite number of mass points. Although in this case we have focused on the MISOSE case, our proof technique can be extended in a straightforward manner to encompass a channel where the number of receive antennas at both the legitimate receiver and the eavesdropper are arbitrary. At high-SNR, it is established that the secret-key capacity is bounded

irrespective of SNR, and regardless of the number of antennas at each terminal. At low-SNR, it has been observed through numerical results that two mass point, one of them at the origin, is optimal.

APPENDIX A
PROOF OF (8)

$$\begin{aligned}
I(X;Y|Z) &= I(X;Y,Z) - I(X;Z) \\
&= h(Y,Z) - h(Y,Z|X) - h(Z) + h(Z|X) \\
&= h(Y|Z) - h(Y|X) \\
&= \iint p_{Y,X}(y,x) \ln\big(p_{Y|X}(y|x)\big) dydx \\
&\quad - \iint p_{Y,Z}(y,z) \ln\big(p_{Y|Z}(y|z)\big) dydz \\
&= \iiint p_{Y,X,Z}(y,x,z) \ln\big(p_{Y|X}(y|x)\big) dydxdz \\
&\quad - \iiint p_{Y,Z|X}(y,z|x) \ln\big(p_{Y|Z}(y|z;F)\big) dydzdF(x) \\
&= \iiint p_{Y,Z|X}(y,z|x) \ln\big(p_{Y|X}(y|x)\big) dydF(x)dz \\
&\quad - \iiint p_{Y|X}(y|x) p_{Z|X}(z|x) \ln\big(p_{Y|Z}(y|z)\big) dydzdF(x)
\end{aligned} \tag{43}$$

where to obtain (43) and (44), we used the fact that given X, Y and Z are independent; and where $p_{Y|Z}(y|z;F)$ signifies the conditional distribution of Y given Z induced by F.

APPENDIX B
PROOF OF LEMMA 4

The probability $p_{Y|Z}(y|z)$ can be written as

$$\begin{aligned}
p_{Y|Z}(y|z;F^*) &= \frac{p_{Y,Z}(y,z;F^*)}{p_Z(z;F^*)} \\
&= \frac{f(y+z)}{g(z)},
\end{aligned} \tag{45}$$

where it follows via (9) that when $n_R = n_E = 1$, the functions $f(\cdot)$ and $g(\cdot)$ in (45) are defined by (26) and (27) respectively.

Thus we can split the left hand side (LHS) of (26) into two parts as stated in (23)

$$\iint e^{-z_c(y+z)} \ln\big[p_{Y|Z}(y|z;F^*)\big] dydz = I_1(z_c) - I_2(z_c),$$

where $I_1(z_c) = \iint e^{-z_c(y+z)} \ln[f(y+z)] dydz$ and $I_2(z_c) = \iint e^{-z_c(y+z)} \ln[g(z)] dydz$. Now, transforming the coordinate system from Y-Z to W-Z where W = Y + Z, we find that the Jacobian of the transformation is 1 and hence I_1 can be written as:

$$I_1(z_c) = \int_0^\infty \int_z^\infty e^{-z_c w} \ln[f(w)] dwdz \tag{46}$$

$$= \int_0^\infty \int_0^w e^{-z_c w} \ln[f(w)] dzdw \tag{47}$$

$$= \int_0^\infty w e^{-z_c w} \ln[f(w)]\, dw \tag{48}$$

On simplification of $I_2(z_c)$, we also get:

$$I_2(z_c) = \frac{1}{z_c} \int_0^\infty e^{-z_c w} \ln[g(w)] dw \tag{49}$$

Note that (48) and (49) represent the Laplace transforms of $w \ln[f(w)]$ and $\int_0^w \ln[g(\tau)] d\tau$, respectively. Since the integrands in (48) and (49) are integrable over the interval $[0,a]$ for all $a > 0$, then to prove that $I_1(z_c)$ and $I_2(z_c)$ are well-defined for all $\Re(z_c) > 0$, it suffices to show that $|we^{-z_c w} \ln[f(w)]|$ and $|e^{-z_c w} \ln[g(w)]|$ are bounded by some integrable functions on $[0,\infty]$. This can be established as follows:

$$\begin{aligned}
|we^{-z_c w} \ln f(w)| &= -we^{-\Re(z_c)w} \ln(f(w)) && (50)\\
&= -we^{-\Re(z_c)w} \ln\big(\mathbb{E}\big[S^2 e^{-Sw}\big]\big) \\
&\le -we^{-\Re(z_c)w} \mathbb{E}\big[\ln\big(S^2 e^{-Sw}\big)\big] && (51)\\
&= -we^{-\Re(z_c)w} \mathbb{E}\big[\ln\big(S^2\big) - Sw\big] \\
&= we^{-\Re(z_c)w} \left(w\mathbb{E}\left[\frac{1}{1+X^2}\right] + 2\mathbb{E}\big[\ln\big(1+X^2\big)\big] \right) \\
&\le we^{-\Re(z_c)w} \left(w + 2\ln(1+P)\right) && (52)\\
&= e^{-\Re(z_c)w} P(w) && (53)
\end{aligned}$$

where in (53), $P(w) = w\left(w + 2\ln(1+P)\right)$ is a polynomial of order 2 in w. (50) holds because $f(w)$ is a decreasing function over $w \ge 0$ and $f(0) < 1$ (c.f. (29)); (51) follows from Jensen inequality; (52) is true because $\frac{1}{1+X^2} \le 1$ and applying Jensen inequality again. Since $\int_0^\infty e^{-\Re(z_c)w} P(w) dw$ exists for all $\Re(z_c) > 0$, then so does $I_1(z_c)$. By a similar technique, the following upper bound may be obtained:

$$|e^{-z_c w} \ln g(w)| \le e^{-\Re(z_c)w} \left(w + \ln(1+P)\right) \tag{54}$$

to justify the convergence of the integral in (49). We conclude that $I_2(z_c)$ is also well-defined for all $\Re(z_c) > 0$.

REFERENCES

[1] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *Information Theory, IEEE Transactions on*, vol. 39, no. 4, pp. 1121 –1132, jul 1993.

[2] U. Maurer, "Secret key agreement by public discussion from common information," *Information Theory, IEEE Transactions on*, vol. 39, no. 3, pp. 733 –742, may 1993.

[3] T. F. Wong, M. Bloch, and J. M. Shea, "Secret sharing over fast-fading mimo wiretap channels," *EURASIP J. Wirel. Commun. Netw.*, vol. 2009, pp. 1–1, 2009.

[4] R. Wilson, D. Tse, and R. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *Information Forensics and Security, IEEE Transactions on*, vol. 2, no. 3, pp. 364 –375, sept. 2007.

[5] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *Information Theory, IEEE Transactions on*, vol. 45, no. 1, pp. 139 –157, jan 1999.

[6] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The Capacity of Discrete-Time Memoryless Rayleigh-Fading Channels," *Information Theory, IEEE Transactions on*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[7] A. Agrawal, Z. Rezki, A. Khisti, and M.-S. Alouini, "Non-coherent capacity of secret-key agreement with public discussion," *submitted to the special issue on IEEE Transactions on Information Forensics and Security*, Sep. 2010.

[8] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2453–2469, 2008. [Online]. Available: http://dx.doi.org/10.1109/TIT.2008.921861

[9] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature gaussian channels," *Information Theory, IEEE Transactions on*, vol. 41, no. 4, pp. 1060 –1071, jul 1995.

[10] H. Silverman, Complex Variables. Boston, MA: Houghton Mifflin, 1975.

[11] Davies Brian , Integral transforms and their applications, 3rd ed. New York: Springer, 2002.