Gaussian Wiretap Channel With Amplitude and Variance Constraints

Omur Ozel, *Member, IEEE*, Ersen Ekrem, and Sennur Ulukus, *Member, IEEE*

***Abstract*— We consider the Gaussian wiretap channel with amplitude and variance constraints on the channel input. We first show that the entire rate-equivocation region of the Gaussian wiretap channel with an amplitude constraint is obtained by discrete input distributions with finite support. We prove this result by considering the existing single-letter description of the rate-equivocation region, and showing that discrete distributions with finite support exhaust this region. Our result highlights an important difference between the peak power (amplitude) constrained and the average power (variance) constrained cases. Although, in the average power constrained case, both the secrecy capacity and the capacity can be achieved simultaneously, our results show that in the peak power constrained case, in general, there is a tradeoff between the secrecy capacity and the capacity, in the sense that, both may not be achieved simultaneously. We also show that under sufficiently small amplitude constraints the possible tradeoff between the secrecy capacity and the capacity does not exist and they are both achieved by the symmetric binary distribution. Finally, we prove the optimality of discrete input distributions in the presence of an additional variance constraint.**

***Index Terms*— Gaussian wiretap channel, rate-equivocation region, amplitude and variance constraints.**

I. Introduction

WE CONSIDER the Gaussian wiretap channel [1]–[3] which consists of a transmitter, a legitimate user and an eavesdropper as shown in Fig. 1. In the Gaussian wiretap channel, each link is a memoryless additive white Gaussian noise (AWGN) channel. In this model, the goal of the transmitter is to have secure communication with the legitimate user while keeping the eavesdropper ignorant of this communication as much as possible.

Since the Gaussian wiretap channel is stochastically degraded, its rate-equivocation region is known in a single-letter form due to [1]. Under an average power constraint, the entire rate-equivocation region of the Gaussian wiretap channel can be obtained by evaluating this single-letter expression. In particular, under an average power constraint, Gaussian input with full power attains both the secrecy capacity and the capacity of the channel between the transmitter and the legitimate user, providing the entire rate-equivocation region [3]. One important implication of this result is that the transmitter and the legitimate user do not compromise from their communication rate in order to maximize the equivocation of their communication at the eavesdropper. In other words, there is no tradeoff between the rate and the equivocation for the average power constrained Gaussian wiretap channel.

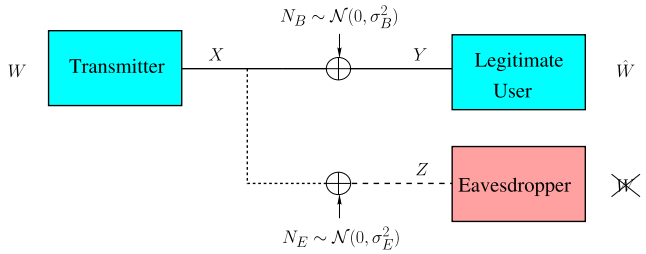

Fig. 1. The Gaussian wiretap channel.

In this work, we consider the Gaussian wiretap channel under amplitude (i.e., peak power) and variance (i.e., average power) constraints. Similar to the average power constrained case, here also, we can use the existing single-letter description for the rate-equivocation region of the Gaussian wiretap channel due to [1]. However, unlike the average power constrained case, here, due to the peak power constraint, the corresponding optimization problems are harder to solve explicitly. For example, the entropy-power inequality, which is the key tool to obtain the rate-equivocation region under an average power constraint [3], does not provide a tight result for the rate-equivocation region under a peak power constraint.

We circumvent difficulties arising from the existence of a peak power constraint by using the methodology originally devised by [4] and [5], and later, extended further by [6]–[13]. In [4] and [5], Smith studied the AWGN channel under peak and average power constraints and proved that the optimal input distribution is discrete with finite support. This methodology considers the functional optimization problem associated with the capacity of the AWGN channel, obtains the necessary and sufficient conditions for the optimal input distribution, and proves by contradiction that the optimal input distribution should be discrete with finite support.

Manuscript received December 27, 2013; revised October 5, 2014; accepted July 12, 2015. Date of publication July 22, 2015; date of current version September 11, 2015. This work was supported by the National Science Foundation under Grant CNS 09-64632, Grant CCF 09-64645, Grant CCF 10-18185, and Grant CNS 11-47811. This paper was presented in part at the IEEE Information Theory Workshop (ITW), Lausanne, Switzerland, September 2012.

O. Ozel was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA. He is now with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: ozel@berkeley.edu).

E. Ekrem was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA. He is now with Qualcomm, Santa Clara, CA 95051 USA (e-mail: ersenek@gmail.com).

S. Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ulukus@umd.edu).

Communicated by T. Liu, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2015.2459705

In this work, we use this methodology [4], [5], [8], [11] to study the Gaussian wiretap channel with amplitude and variance constraints. First, we consider the single-letter description of the rate-equivocation region under a peak power constraint, and obtain necessary and sufficient conditions for the optimal input distribution. Next, we prove by contradiction that the optimal input distribution should be discrete with finite support. We provide numerical results which highlight an important difference between the peak power constrained and the average power constrained cases. As mentioned, in the average power constrained case, both the secrecy capacity and the capacity are simultaneously achieved by the same input distribution (Gaussian distribution with full power). On the other hand, our numerical results demonstrate that under a peak power constraint, in general, the secrecy capacity and the capacity are not achieved by the same distribution. In other words, under a peak power constraint, in general, there is a tradeoff between the rate and its equivocation, in the sense that, when we want to maximize the equivocation, we may need to compromise from the rate; and conversely, when we want to maximize the rate, we may need to compromise from its equivocation.

Next, we study the conditions under which a binary input distribution is optimal in the amplitude constrained Gaussian wiretap channel. By adapting the steps in [12] for the Gaussian wiretap channel, we show that if $A \leq 1.05$, the rate-equivocation region boundary is achieved by the symmetric binary distribution. In other words, there is no tradeoff between the rate and its equivocation if the amplitude constraint is sufficiently small.

Finally, we extend the optimality of discrete input distributions to the case when an additional variance constraint is imposed on the input. To this end, we provide a modified contradiction argument that uses the optimality conditions of the equivalent amplitude unconstrained optimization problem. In particular, we start with the KKT optimality conditions of the amplitude and variance constrained problem and show, using analyticity and the identity theorem, that these KKT conditions are equivalent to the KKT conditions of the amplitude unconstrained and variance constrained problem. The unique solution of the amplitude unconstrained and variance constrained problem is known to be a Gaussian distribution. Since the Gaussian distribution is not amplitude constrained, this yields a contradiction. We present this modified contradiction argument in Appendix B for the single-user AWGN channel, and adapt it in the text for the Gaussian wiretap channel.

II. Amplitude Constrained Gaussian Wiretap Channel

The Gaussian wiretap channel is defined by

$$Y_i = X_i + N_{B_i}, \quad i = 1, \ldots, n \tag{1}$$

$$Z_i = X_i + N_{E_i}, \quad i = 1, \ldots, n \tag{2}$$

where X_i, Y_i, Z_i denote the channel input, the legitimate user's observation and the eavesdropper's observation, respectively. N_{B_i} and N_{E_i} are i.i.d. zero-mean Gaussian random variables with variances σ_B^2 and σ_E^2, respectively, where $\sigma_B^2 < \sigma_E^2$. We assume that there is an amplitude constraint on the channel input X_i as

$$|X_i| \leq A, \quad i = 1, \ldots, n \tag{3}$$

An $(n, 2^{nR})$ code for the Gaussian wiretap channel with peak power constraint consists of a message set $W \in \mathcal{W} = \{1, \ldots, 2^{nR}\}$, an encoder at the transmitter $f_n : \mathcal{W} \to \mathbb{R}^n$ satisfying the peak power constraint in (3), and a decoder at the legitimate user $g_n : \mathbb{R}^n \to \mathcal{W}$. Equivocation of a code is measured by the normalized conditional entropy $(1/n)H(W|Z^n)$, where W is a uniformly distributed random variable over $\mathcal{W}$. Probability of error for a code is defined as $P_e^n = \Pr[g_n(f_n(W)) \neq W]$. A rate-equivocation pair (R, R_e) is said to be achievable if there exists an $(n, 2^{nR})$ code satisfying $\lim_{n\to\infty} P_e^n = 0$, and

$$R_e \leq \lim_{n\to\infty} \frac{1}{n} H(W|Z^n) \tag{4}$$

The rate-equivocation region consists of all achievable rate-equivocation pairs, and is denoted by $\mathcal{C}$. A rate R is said to be perfectly secure if we have $R_e = R$, i.e., if there exists an $(n, 2^{nR})$ code satisfying $\lim_{n\to\infty}(1/n)I(W; Z^n) = 0$. Supremum of such rates is defined to be the secrecy capacity and denoted by C_s.

Since the Gaussian wiretap channel is stochastically degraded, its entire rate-equivocation region $\mathcal{C}$ can be expressed in a single-letter form by using the result of [1].

Theorem 1: The rate-equivocation region of the Gaussian wiretap channel with a peak power constraint is given by the union of the rate-equivocation pairs (R, R_e) satisfying

$$R \leq I(X; Y) \tag{5}$$

$$R_e \leq I(X; Y) - I(X; Z) \tag{6}$$

for some input distribution $F_X \in \mathbf{\Omega}$, where the feasible set $\mathbf{\Omega}$ is given by

$$\mathbf{\Omega} \triangleq \left\{ F_X : \int_{-A}^{A} dF_X(x) = 1 \right\} \tag{7}$$

Since the rate-equivocation region $\mathcal{C}$ is convex due to time-sharing, it can be characterized by finding the tangent lines to the region $\mathcal{C}$, which are given by the solutions of

$$\max_{F_X \in \mathbf{\Omega}} g_\mu(F_X) = \max_{F_X \in \mathbf{\Omega}} \mu I(X; Y) + I(X; Y) - I(X; Z) \tag{8}$$

for all $\mu \geq 0$.

For the amplitude constrained Gaussian wiretap channel, our main result is to show that the maximizer distribution for (8) is discrete with finite support.

Theorem 2: Let F_X^* be the maximizer of the optimization problem in (8) with a support set $\mathcal{S}_{F_X^*}$. The support set $\mathcal{S}_{F_X^*}$ is a finite set.

Theorem 2 implies that the secrecy capacity C_s is also achieved by a discrete distribution with finite support, as stated in the following corollary.

Corollary 1: Let F_X^* be the distribution that attains the secrecy capacity of the Gaussian wiretap channel with a peak power constraint. The support set $\mathcal{S}_{F_X^*}$ is a finite set.

In the next two subsections, we first prove Corollary 1, and then, by using the proof of Corollary 1, we prove Theorem 2.

A. Proof of Corollary 1

The proof follows from the convexity of the optimization problem [14] and hence the fact that derivation of an equivalent necessary and sufficient optimality condition in terms of equivocation density is possible [4], [5]. Then, we provide a contradiction argument to prove that a support set with infinite points cannot be optimal under an amplitude constrained input. We start by noting that the secrecy capacity of the Gaussian wiretap channel with peak power constraint is given by

$$C_s = \max_{F_X \in \Omega} g_0(F_X) = \max_{F_X \in \Omega} I(X;Y) - I(X;Z) \tag{9}$$

where the objective function $g_0(F_X)$ is a strictly concave functional of the input distribution F_X due to the assumption $\sigma_B^2 < \sigma_E^2$ [14]. Moreover, the feasible set Ω is convex and sequentially compact with respect to the Levy metric [4]. Thus, (9) is a convex optimization problem with a unique solution.

Next, we obtain the necessary and sufficient conditions that the optimal distribution F_X^* of the optimization problem in (9) should satisfy. To this end, we introduce some notation which will be frequently used throughout the paper. Since both channels are AWGN, the output densities for Y and Z exist for any input distribution F_X, and are given by

$$p_Y(y;F_X) = \int_{-A}^{A} \phi_B(y-x) dF_X(x), \quad y \in \mathbb{R} \tag{10}$$

$$p_Z(z;F_X) = \int_{-A}^{A} \phi_E(z-x) dF_X(x), \quad z \in \mathbb{R} \tag{11}$$

where $\phi_B(y), \phi_E(z)$ are zero-mean Gaussian densities with variances σ_B^2 and σ_E^2, respectively.

We define the equivocation density $r_e(x;F_X)$ as

$$r_e(x;F_X) = i_B(x;F_X) - i_E(x;F_X) \tag{12}$$

where $i_B(x;F_X)$ and $i_E(x;F_X)$ are the mutual information densities for the main channel and the wiretapper's channel

$$i_B(x;F_X) = -\phi_B(x) * \log\left(p_Y(x;F_X)\right) - \frac{1}{2}\log\left(2\pi e \sigma_B^2\right) \tag{13}$$

$$i_E(x;F_X) = -\phi_E(x) * \log\left(p_Z(x;F_X)\right) - \frac{1}{2}\log\left(2\pi e \sigma_E^2\right) \tag{14}$$

where $*$ denotes the convolution. We note that the convolutions in (13) and (14) follow from the symmetry of the Gaussian density function. The mutual information and the mutual information density are related through

$$I(X;Y) = \int_{-A}^{A} i_B(x;F_X) dF_X(x) \tag{15}$$

$$I(X;Z) = \int_{-A}^{A} i_E(x;F_X) dF_X(x) \tag{16}$$

Since the Gaussian wiretap channel is stochastically degraded, without loss of generality, we can assume $Z = Y + Z_D$ for some zero-mean Gaussian random variable Z_D with variance $\sigma_D^2 = \sigma_E^2 - \sigma_B^2$. We denote the density of Z_D by $\phi_D(x)$ which leads to the identity $\phi_E = \phi_B * \phi_D$. Using this identity in conjunction with (13)-(14), the equivocation density $r_e(x;F_X)$ in (12) can be expressed as

$$r_e(x;F_X) = \frac{1}{2}\log\left(\frac{\sigma_E^2}{\sigma_B^2}\right) - \phi_B(x) * \Big[\log\left(p_Y(x;F_X)\right) - \phi_D(x) * \log\left(p_Z(x;F_X)\right)\Big] \tag{17}$$

Now, we are ready to obtain the necessary and sufficient conditions for the optimal distribution of the optimization problem in (9). To this end, we first note that the objective function $g_0(F_X)$ in (9) is Frechet differentiable and the derivative of $g_0(F_X)$ at F_{X_0} in the direction of F_X is given by:

$$\begin{aligned}\lim_{\theta\to 0}\frac{1}{\theta}\left[g_0(\theta F_X + (1-\theta)F_{X_0}) - g_0(F_{X_0})\right] &= \int_{\mathbb{R}} \left(p_Y(y;F_{X_0}) - p_Y(y;F_X)\right)\log\left(p_Y(y;F_{X_0})\right) dy \\ &\quad - \int_{\mathbb{R}} \left(p_Z(z;F_{X_0}) - p_Z(z;F_X)\right)\log\left(p_Z(z;F_{X_0})\right) dz\end{aligned} \tag{18}$$

which, using the equivocation density in (17), is expressed as

$$\lim_{\theta\to 0}\frac{1}{\theta}\left[g_0(\theta F_{X_0} + (1-\theta)F_X) - g_0(F_X)\right] = \int_{-A}^{A} r_e(x;F_{X_0}) dF_X(x) - g_0(F_{X_0}) \tag{19}$$

Due to the linearity of the derivative operation, the Frechet derivative of $g_0(F_X)$ in (18) is the difference of Frechet derivatives of $I(X;Y)$ and $I(X;Z)$. Explicit derivations of the Frechet derivatives of individual mutual information terms can be found in [4, Proof of Proposition 1] and [5, Lemma on p. 29].

In view of the concavity of the objective functional in (9) with respect to the input distribution F_X, steps analogous to [4, Corollary 1] yield the following necessary and sufficient conditions for the optimality of the distribution F_X^*:

$$r_e(x;F_X^*) \le C_s, \quad \forall x \in [-A, A] \tag{20}$$

$$r_e(x;F_X^*) = C_s, \quad \forall x \in \mathcal{S}_{F_X^*} \tag{21}$$

where the secrecy capacity C_s is expressed as

$$C_s = I_B(F_X^*) - I_E(F_X^*) = h_Y(F_X^*) - h_Z(F_X^*) + \frac{1}{2}\log\left(\frac{\sigma_E^2}{\sigma_B^2}\right) \tag{22}$$

where $I_B(F_X^*)$ and $I_E(F_X^*)$ are the mutual information for Bob (between X and Y) and Eve (between X and Z), respectively, generated by the input distribution F_X^*. Similarly, $h_Y(F_X^*)$ and $h_Z(F_X^*)$ are the differential entropies of Y and Z, respectively, generated by the input distribution F_X^*. We note that (20)-(21) are equivalent to the Kuhn-Tucker conditions for the functional optimization problem in (9). Due to the

concavity of the objective in (9), non-negativity of the Frecehet derivative in (18) in every direction is necessary and sufficient, see [4, Proposition 1]. This, in turn, is equivalent to (20)-(21) by [4, Corollary 1].

We now prove by contradiction that the support set $\mathcal{S}_{F_X^*}$ of the optimal distribution is a finite set. To reach a contradiction, we use the optimality conditions given by (20)-(21). To this end, we note that both $i_B(x; F_X)$ and $i_E(x; F_X)$ have analytic extensions over the whole complex plane $\mathbb{C}$ [4]. Since $\phi_B(z)$, $\phi_E(z)$ have analytic extensions for all $z \in \mathbb{C}$, the following functions of a complex variable are well defined and analytic for all $z \in \mathbb{C}$:

$$i_B(z; F_X) = -\int_{-\infty}^{\infty} \phi_B(z-\tau) \log\left(p_Y(\tau; F_X)\right) d\tau - \frac{1}{2}\log\left(2\pi e \sigma_B^2\right) \tag{23}$$

$$i_E(z; F_X) = -\int_{-\infty}^{\infty} \phi_E(z-\tau) \log\left(p_Z(\tau; F_X)\right) d\tau - \frac{1}{2}\log\left(2\pi e \sigma_E^2\right) \tag{24}$$

Therefore, the equivocation density has the analytic extension $r_e(z; F_X) = i_B(z; F_X) - i_E(z; F_X)$ for $z \in \mathbb{C}$. Now, let us assume that $\mathcal{S}_{F_X^*}$ has infinite number of elements. In view of the optimality condition (21), analyticity of $r_e(z; F_X)$ over all $\mathbb{C}$ and the identity theorem for complex numbers along with Bolzano-Weierstrass theorem, if $\mathcal{S}_{F_X^*}$ has infinite number of elements, we should have $r_e(z; F_X^*) = C_s$ for all $z \in \mathbb{C}$, which, in turn, implies

$$r_e(x; F_X^*) = C_s, \quad \forall x \in \mathbb{R} \tag{25}$$

Next, we show that (25) results in a contradiction. To this end, we first state the following result from [11].

Lemma 1 ([11, Corollary 9]): Let Z be a Gaussian random variable and $P_Z(z)$ be the corresponding probability density function. Suppose $g(z)$ is a continuous function such that $|g(z)| \leq \alpha + \beta|z|^2$ for some $\alpha, \beta > 0$. If $P_Z(z) * g(z)$ is the zero function, then $g(z)$ is also the zero function.

Next, we rearrange (25) by using (17) to get

$$\int_{\mathbb{R}} \phi_B(y-x) v(y) dy = 0, \quad \forall x \in \mathbb{R} \tag{26}$$

where $v(y)$ and c are defined as

$$v(y) = c + \log\left(p_Y(y; F_X^*)\right) - \int_{\mathbb{R}} \phi_D(\tau) \log\left(p_Z(y-\tau; F_X^*)\right) d\tau \tag{27}$$

$$c = h_Y(F_X^*) - h_Z(F_X^*) \tag{28}$$

Note that $c < 0$ for any nontrivial input distribution F_X^*. This follows from the stochastic degradedness of the channel, i.e., $Z = Y + Z_D$ for some zero-mean Gaussian random variable Z_D with variance $\sigma_D^2 = \sigma_E^2 - \sigma_B^2$. Hence $h(Z) > h(Z|Z_D) = h(Y)$ by the fact that conditioning reduces entropy, and this proves $c < 0$. Note that $h_Y(F_X^*)$ and $h_Z(F_X^*)$ are representations of $h(Y)$ and $h(Z)$, respectively.

Next, we show that if (26) holds, we should have $v(y) = 0, \forall y \in \mathbb{R}$. To this end, we note that since $p_Y(y; F_X^*) = \int_{-A}^{A} \phi_B(y-x) dF_X^*(x)$, Jensen's inequality implies

$$\frac{1}{\sqrt{2\pi\sigma_B^2}} \geq p_Y(y; F_X^*) \geq \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-\frac{1}{2\sigma_B^2}\int_{-A}^{A}(y-x)^2 dF_X^*(x)} \tag{29}$$

which, in turn, implies $|\log\left(p_Y(y; F_X^*)\right)| \leq \alpha y^2 + \beta$ for some $\alpha, \beta > 0$. Similarly, we can show that $|\log\left(p_Z(y; F_X^*)\right)| \leq \kappa y^2 + \gamma$ for some $\kappa, \gamma > 0$. Consequently, we have $|v(y)| \leq \eta y^2 + \zeta$ for some $\eta, \zeta > 0$, which, in conjunction with (26) and by Lemma 1, implies that $v(y) = 0$ for all $y \in \mathbb{R}$.

Now, we show that we cannot have $v(y) = 0, \forall y \in \mathbb{R}$, and therefore, reach a contradiction. In particular, we show that there exists y' such that $v(y) < 0, \forall y \geq y'$. To this end, we note that $c < 0$ and introduce the following lemma.

Lemma 2: There exists sufficiently large y' such that $\forall y \geq y'$, we have

$$\int_{\mathbb{R}} \phi_D(\tau) \log\left(p_Z(y-\tau; F_X^*)\right) d\tau \geq \log\left(p_Y(y; F_X^*)\right) \tag{30}$$

We provide the proof of Lemma 2 in Appendix A.

Lemma 2 and the fact that $c < 0$ imply that $v(y) < 0$, $\forall y \geq y'$, which, in turn, implies that (26) cannot hold. This, in turn, implies that $\mathcal{S}_{F_X^*}$ cannot have infinite number of elements. This completes the proof of Corollary 1.

B. Proof of Theorem 2

In this section, we extend our analysis in the previous section to the entire rate-equivocation region. This extension entails generalizing the contradiction argument in the proof of Corollary 1 to the case when an additional mutual information term is present in the objective function. We start by noting that the rate-equivocation region can be characterized by solving the following optimization problem

$$\max_{\mathcal{F}_X \in \Omega} g_\mu(F_X) = \max_{\mathcal{F}_X \in \Omega} \mu I(X;Y) + I(X;Y) - I(X;Z) \tag{31}$$

for all $\mu \geq 0$. Since the objective function $g_\mu(F_X)$ in (31) is strictly concave, and the feasible set Ω is convex and sequentially compact with respect to the Levy metric, the optimization problem in (31) has a unique maximizer. We denote the optimal input distribution for (31) as F_X^* which depends on the value of μ.

Now, we obtain the necessary and sufficient conditions for the optimal distribution of the optimization problem in (31). To this end, we note that $g_\mu(F_X)$ is Frechet differentiable, and its derivative at F_{X_0} in the direction of F_X is given as

$$\lim_{\theta \to 0} \frac{1}{\theta}\left[g_\mu(\theta F_X + (1-\theta)F_{X_0}) - g_\mu(F_{X_0})\right] = \int_{-A}^{A} \left[\mu i_B(x; F_{X_0}) + r_e(x; F_{X_0})\right] dF_X(x) - g_\mu(F_{X_0}) \tag{32}$$

Using similar arguments to those in [4], the necessary and sufficient conditions for the optimal distribution of the

optimization problem in (31) can be obtained as follows

$$\mu i_B(x;F_X^*)+r_e(x;F_X^*)\leq(\mu+1)I_B(F_X^*)-I_E(F_X^*),\quad \forall x\in[-A,A] \tag{33}$$

$$\mu i_B(x;F_X^*)+r_e(x;F_X^*)=(\mu+1)I_B(F_X^*)-I_E(F_X^*),\quad \forall x\in\mathcal{S}_{F_X^*} \tag{34}$$

where $I_B(F_X^*)$ and $I_E(F_X^*)$ are the mutual information for Bob (between X and Y) and Eve (between X and Z), respectively, generated by the input distribution F_X^*. Similarly, $i_B(x;F_X^*)$ and $i_E(x;F_X^*)$ are the corresponding mutual information densities generated by F_X^*.

Now, we show that the optimal input distribution F_X^* should have finite support. Similar to the proof of Corollary 1, here also, we prove the finiteness of the support set by contradiction and using the optimality conditions in (33)-(34).

Let us assume that $\mathcal{S}_{F_X^*}$ has infinite number of elements. Under this assumption, (34), analyticity of $i_B(z;F_X^*)$ and $r_e(z;F_X^*)$ over all $\mathbb{C}$ and the identity theorem for complex numbers imply that $\mu i_B(z;F_X^*)+r_e(z;F_X^*)=(\mu+1)I_B(F_X^*)-I_E(F_X^*)$ over all $\mathbb{C}$, which, in turn, implies that

$$\mu i_B(x;F_X^*)+r_e(x;F_X^*)=(\mu+1)I_B(F_X^*)-I_E(F_X^*),\quad \forall x\in\mathbb{R} \tag{35}$$

Next, we show that (35) results in a contradiction. To this end, we first rearrange (35) to obtain

$$\int_{\mathbb{R}}\phi_B(y-x)\hat{v}(y)dy=0 \tag{36}$$

where $\hat{v}(y)$ and $\hat{c}$ are given by

$$\hat{v}(y)=\hat{c}+(\mu+1)\log\left(p_Y(y;F_X^*)\right)-\int_{\mathbb{R}}\phi_D(\tau)\log\left(p_Z(y-\tau;F_X^*)\right)d\tau \tag{37}$$

$$\hat{c}=(\mu+1)h_Y(F_X^*)-h_Z(F_X^*) \tag{38}$$

We note that the expressions in (37)-(38) differ from the ones in (27)-(28) for the secrecy capacity in the additional terms factored by μ; hence, the negativity of $\hat{c}$ is not immediately ensured. Therefore, we take an alternative route for the proof. By using similar arguments to those we provided in the proof of Corollary 1, one can show that $|\hat{v}(y)|\leq\eta y^2+\zeta$ for some $\eta,\zeta>0$. By Lemma 1, this implies that if (36) holds, we should have $\hat{v}(y)=0,\forall y\in\mathbb{R}$. Next, we show that we cannot have $\hat{v}(y)=0,\forall y\in\mathbb{R}$. Using Lemma 2 and the fact that $h_Y(F_X^*)-h_Z(F_X^*)<0$ in (37), we get

$$\hat{v}(y)-\mu\left(h_Y(F_X^*)+\log\left(p_Y(y;F_X^*)\right)\right)<0,\quad \forall y\geq y' \tag{39}$$

Hence, if $\hat{v}(y)=0,\forall y\in\mathbb{R}$ holds, due to (39), we should have

$$h_Y(F_X^*)+\log\left(p_Y(y;F_X^*)\right)>0,\quad \forall y\geq y' \tag{40}$$

which implies

$$p_Y(y;F_X^*)\geq e^{-h_Y(F_X^*)},\quad \forall y\geq y' \tag{41}$$

However, since $p_Y(y;F_X^*)$ is a density function, it has to vanish as $y\to\infty$, and (41) cannot hold. Hence, we reach a contradiction; implying that the optimal input distribution should have a finite support set. This completes the proof of Theorem 2.

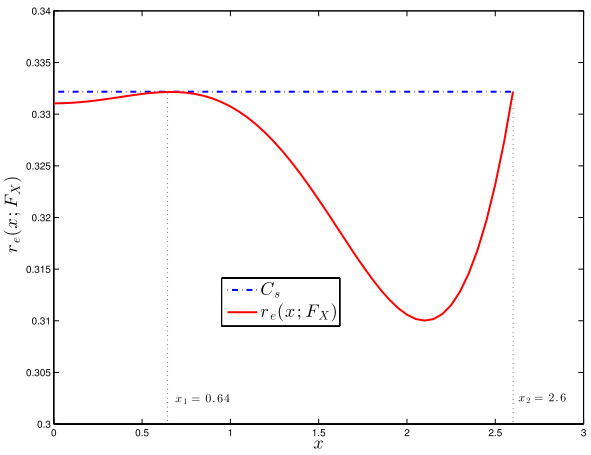

Fig. 2. Illustration of the equivocation density yielded by the optimal input distribution when $\sigma_B^2=1$, $\sigma_E^2=2$ and $A=2.6$. Note that the equivocation density is symmetric around $x=0$ and the plot is shown only for $x\geq 0$.

III. Numerical Results for the Amplitude Constrained Case

In this section, we provide numerical illustrations for the secrecy capacity and the rate-equivocation region of the Gaussian wiretap channel under a peak power constraint.

We first consider how the secrecy capacity changes with respect to the amplitude constraint A for $\sigma_B^2=1$ and $\sigma_E^2=2$. We provide a plot of the equivocation density for an optimal input distribution in Fig. 2 for $A=2.6$. We numerically calculated that for these parameters the optimal input distribution is quaternary located at $x=\pm 0.64$ and $x=\pm 2.6$ with probability masses 0.2496 at $x=\pm 0.64$ and 0.2504 at $x=\pm 2.6$. We observe that the equivocation density is less than or equal to the secrecy capacity and it is equal to the secrecy capacity at the mass points; verifying the optimality conditions in (20)-(21).

Next, we observe in Fig. 3 that the rates of increase of the amplitude and variance constrained capacities with respect to SNR follow the same asymptote. A similar observation was made by Smith [4] for the capacities without secrecy constraint. Moreover, in Fig. 3, we also plot the difference C_B-C_E where C_B and C_E are the legitimate user's and the eavesdropper's capacities, respectively. This difference is, in general, a lower bound for the secrecy capacity C_s. We observe that, for small values of A, C_B-C_E and C_s are identical.[1] However, as A increases, C_B-C_E and C_s become different. We note that $I(X;Y)$, $I(X;Z)$ and the difference $I(X;Y)-I(X;Z)$ are concave in the input distribution. Hence, one may be tempted to conclude that if the same input distribution maximizes both $I(X;Y)$ and $I(X;Z)$, then it should also maximize the difference $I(X;Y)-I(X;Z)$. However, this observation holds if the capacity achieving input distribution is within the interior of the feasible set; but not on the boundary, see also [14, Th. 3]. For the average power constrained case, Gaussian distribution maximizes both $I(X;Y)$

[1]We will investigate this analytically in Section IV.

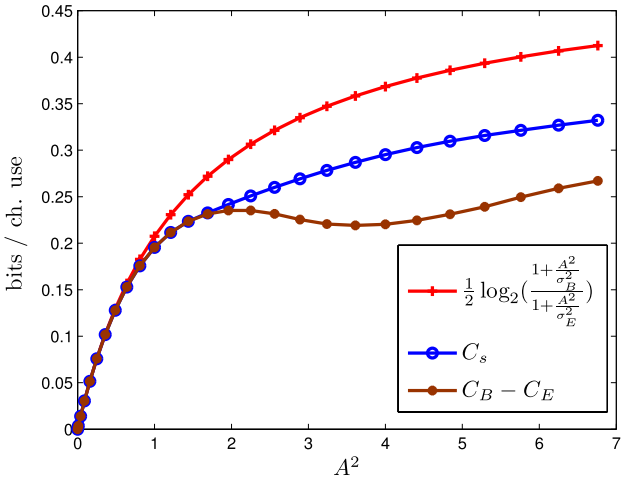

Fig. 3. The secrecy capacity for $\sigma_B^2 = 1$ and $\sigma_E^2 = 2$ versus the square of the amplitude constraint A.

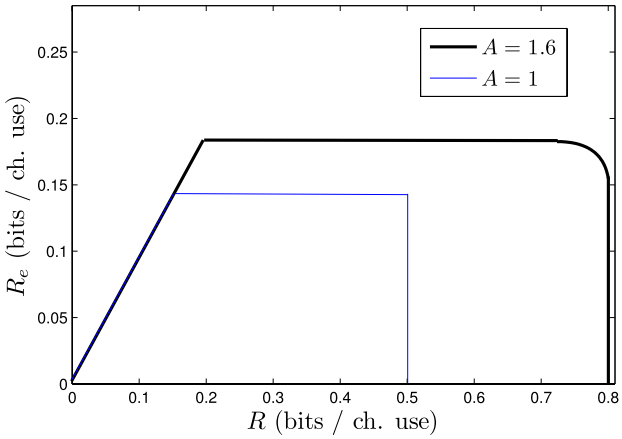

Fig. 4. The rate-equivocation regions for $\sigma_B^2 = 1$ and $\sigma_E^2 = 1.6$ under amplitude constraints $A = 1$ and $A = 1.6$.

and $I(X;Z)$ and as the Gaussian distribution is not on the boundary of the feasible set, it also maximizes the difference $I(X;Y) - I(X;Z)$. However, for the peak power constrained case, discrete distributions are extreme distributions, lying out of the interior of the space of input distributions. Therefore, even if both $I(X;Y)$ and $I(X;Z)$ are maximized by the same discrete distribution, $I(X;Y) - I(X;Z)$ may be maximized by a different input distribution. As a specific example, when $A = 1.5$ and hence $A^2 = 2.25$, while both $I(X;Y)$ and $I(X;Z)$ are maximized by the same binary distribution with equal probability masses at $\pm A$, $I(X;Y) - I(X;Z)$ is maximized by a ternary distribution with mass points at $\pm A$ and 0 with probabilities 0.399, 0.399 and 0.202, respectively. This explains the difference between C_s and $C_B - C_E$ at $A^2 = 2.25$ in Fig. 3.

In Fig. 4, we plot the entire rate-equivocation region of the wiretap channel when $\sigma_B^2 = 1$ and $\sigma_E^2 = 1.6$ for two different values of A. When $A = 1$, it is clear from Fig. 4 that both the secrecy capacity and the capacity can be attained simultaneously. In particular, for $A = 1$, the binary input distribution located at $\pm A$ with equal probabilities achieves both the capacity and the secrecy capacity. In fact, for $A = 1$, the binary distribution at $\pm A$ with equal probabilities maximizes $I(X;Y)$, $I(X;Z)$ and $I(X;Y) - I(X;Z)$. That is, the optimal input distributions for the secrecy capacity and the capacity are identical. This implies that, when $A = 1$, the transmitter can communicate with the legitimate user at the capacity while achieving the maximum equivocation at the same time. On the other hand, when $A = 1.6$, the secrecy capacity and the capacity cannot be achieved simultaneously. In particular, for $A = 1.6$, the binary input distribution located at $\pm A$ with equal probabilities achieves the capacity, while a ternary distribution located at $x = \pm A$ and $x = 0$ with probability masses 0.358 at $\pm A$ and 0.284 at 0 achieves the secrecy capacity, i.e., the optimal input distributions for the secrecy capacity and the capacity are different. In other words, there is a tradeoff between the rate and the equivocation in the sense that, to increase the communication rate, we should compromise from the equivocation of this communication, and to increase the achieved equivocation, we should compromise from the communication rate. This result is in contrast with the average power constrained case, where irrespective of the average power constraint, both the secrecy capacity and the capacity can be simultaneously achieved by a Gaussian distribution with full power.

IV. On the Optimality of Symmetric Binary Distribution

We have seen that in the peak power constrained case, there may be a tradeoff between the secrecy capacity and the capacity. However, the numerical results in Section III indicate that if the amplitude constraint is sufficiently small, binary distribution achieves both the secrecy capacity and the capacity simultaneously. In this section, we will quantify this observation by extending the result in [12] to the wiretap channel setting.

We first note that the optimal input distributions that solve (8) are always symmetric around the origin.

Lemma 3: The solution of (8) is symmetric around the origin.

Proof: The claim follows from the fact that the Gaussian density is symmetric around the origin and since both channels are additive noise channels, flipping the input distribution around the origin yields the same mutual informations and secrecy rate. Moreover, the objective $g_\mu(F_X) = (\mu + 1) I_{F_X}(X;Y) - I_{F_X}(X;Z)$ is strictly concave with the input distribution. By [15, Proposition 1] (see also [5, Lemma on p. 44]), we get the desired result. ∎

Moreover, there are always non-zero probability mass points at $-A$ and $+A$ when $\mu = \infty$, i.e., when the objective function is $I(X;Y)$; see also [15]. A possible proof for this follows from the I-MMSE relation [16], [17], since $I(X;Y)$ is monotone increasing function of the snr. Therefore, if the amplitude constraint is not satisfied with equality, there is always room for improvement. On the other hand, it is not clear that the mutual information difference $I(X;Y) - I(X;Z)$ is always monotone increasing with the snr and hence the inclusion of $+A$ and $-A$ in the support set of the optimal

input distribution for all $\mu > 0$ is inconclusive. However, we observed in our numerical studies that $+A$ and $-A$ points are always included. A mathematical proof for this remains an open problem.

Next, we will follow steps analogous to those in [12]. We first note that by using the I-MMSE relation in [16] and [17], when $\sigma_B^2 = 1$ we can express the mutual information difference as:

$$I(X;Y) - I(X;Z) = \int_{\frac{1}{\sigma_E^2}}^{1} \text{mmse}(X|\sqrt{\gamma}X + N)d\gamma \quad (42)$$

where $\text{mmse}(X|\sqrt{\gamma}X+N)$ is the minimum mean squared error for the input X given the noisy observation $\sqrt{\gamma}X + N$ where N is a zero-mean unit-variance Gaussian noise independent of X. Note that $\text{mmse}(X|\sqrt{\gamma}X + N)$ is a functional of the input distribution F_X. In [18], it is shown that the least favorable (MMSE maximizing) input distribution is the symmetric binary distribution $\frac{1}{2}\delta_{-A} + \frac{1}{2}\delta_A$ if $|X| \leq A \leq 1.05$ and $\gamma \leq 1$. Therefore, as in [12], the integrand on the right hand side of (42) is always maximized by this binary input distribution for the range $\gamma \in (\frac{1}{\sigma_E^2}, 1)$. This implies that $I(X;Y)$ and $I(X;Y) - I(X;Z)$ are both maximized by the symmetric binary distribution located at $\pm A$ if $A \leq 1.05$.

Theorem 3: If $A \leq 1.05$, the entire rate-equivocation region boundary is achieved by the symmetric binary input distribution $\frac{1}{2}\delta_{-A} + \frac{1}{2}\delta_A$.

Theorem 3 implies that for sufficiently small amplitude constraints, binary distribution is optimal for the secrecy capacity. As we increase the amplitude constraint, optimal distribution changes. Let A_c be the critical maximum amplitude constraint for which the secrecy capacity achieving input distribution is binary. One can numerically calculate A_c for specified $\sigma_B^2 = 1$ and $\sigma_E^2 > 1$ values. In [15], $A = 1.67$ is calculated as the maximum amplitude constraint for which the legitimate user's capacity is achieved by the binary distribution. Accordingly, as $\sigma_E^2 \to \infty$, A_c approaches 1.67. On the other extreme, as $\sigma_E^2 \to 1$, the critical amplitude constraint approaches $A_c = 1.05$ due to the relation in (42) and the fact that the MMSE maximizer distribution transitions from binary to ternary at $A = 1.05$ as calculated in [18]. In Fig. 5, we plot A_c with respect to σ_E^2. The range of A_c is [1.05, 1.67]. A_c monotonically increases from 1.05 to 1.67 as the noise variance of the eavesdropper σ_E^2 increases from 1 to ∞, when $\sigma_B^2 = 1$.

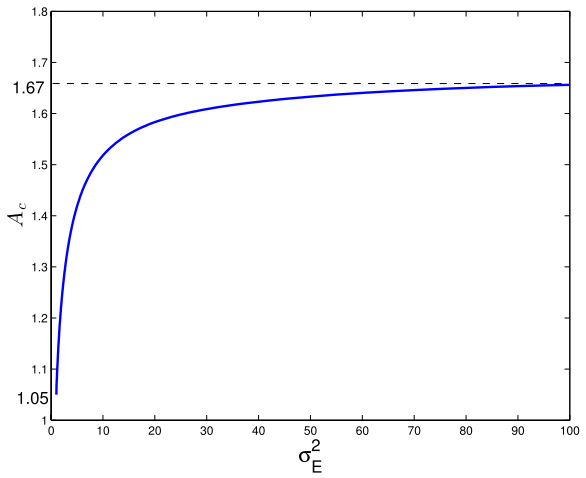

Fig. 5. Critical amplitude A_c where the optimal distribution switches from binary to ternary with respect to σ_E^2.

V. The Case of Amplitude and Variance Constraints

In this section, we generalize the discreteness result for the optimal input distribution when a variance constraint is present in addition to an amplitude constraint. In the AWGN channel with amplitude and variance constraints, the proof of discreteness follows from the fact that if the optimal input distribution F_X^* has infinitely many points in its support set, then it has to be a Gaussian distribution, contradicting the fact that the input is amplitude constrained [4], [5], [11]. This fact is proved in [4], [5], and [11] by the observation (after using properties of Schwartz functions) that the output density $p_Y(y; F_X^*)$ has to be Gaussian distributed if the support set of F_X^* has infinitely many points. In the Gaussian wiretap channel setting under amplitude and variance constraints, proving that $p_Y(y; F_X^*)$ and $p_Z(z; F_X^*)$ have to be both Gaussian distributed if the support of the optimal input distribution F_X^* has infinitely many elements is not straightforward using the properties of Schwartz functions. Therefore, we need an alternative approach to prove the fact that the optimal input distribution is still discrete under amplitude and variance constraints in the Gaussian wiretap channel. To this end, we devise in Appendix B a modified argument for the discreteness proof in [4], [5], and [11] for the AWGN channel with amplitude and variance constraints. In the following, we show that this modified argument is more suitable for our purposes as it easily generalizes to the wiretap channel with amplitude and variance constraints.

We now generalize Theorem 2 and Corollary 1 for the case when we have both amplitude and variance constraints by establishing parallels to the modified proof method presented in Appendix B. Let the variance constraint be P. The new feasible set for the input distribution is

$$\Omega_{A,P} = \left\{ F_X : \int_{-A}^{A} dF_X(x) = 1, \quad \int_{-A}^{A} x^2 dF_X(x) \leq P \right\} \quad (43)$$

We start with considering the secrecy capacity, C_s:

$$C_s = \max_{F_X \in \Omega_{A,P}} I(X;Y) - I(X;Z) \quad (44)$$

In view of [4, Proposition 3] and the strict concavity of the mutual information difference $I(X;Y) - I(X;Z)$ along with the compactness of $\Omega_{A,P}$, the necessary and sufficient conditions in (20)-(21) take the following new form

$$r_e(x; F_X^*) - \lambda x^2 \leq C_s - \lambda E[X^2], \quad \forall x \in [-A, A] \quad (45)$$

$$r_e(x; F_X^*) - \lambda x^2 = C_s - \lambda E[X^2], \quad \forall x \in \mathcal{S}_{F_X^*} \quad (46)$$

$$\lambda \left(E[X^2] - P \right) = 0 \quad (47)$$

for some $\lambda \geq 0$. We note that if the variance constraint is not tight for the optimal distribution F_X^*, then $\lambda = 0$. In this case, F_X^* is the optimal distribution under the amplitude constraint only, which has already been proven in Corollary 1 to be discrete with finite support. Therefore, we assume, without loss of generality, that $\lambda > 0$ and (45)-(47) reduce to:

$$r_e(x; F_X^*) - \lambda x^2 \leq C_s - \lambda P, \quad \forall x \in [-A, A] \tag{48}$$

$$r_e(x; F_X^*) - \lambda x^2 = C_s - \lambda P, \quad \forall x \in \mathcal{S}_{F_X^*} \tag{49}$$

$$E[X^2] = P \tag{50}$$

Next, we prove by contradiction that the input distribution F_X^* that satisfies (48)-(50) must be a discrete distribution with finite support. Assume that $\mathcal{S}_{F_X^*}$ has infinite number of elements. In view of (48)-(50), analyticity of $r_e(z; F_X)$ and z^2 over $\mathbb{C}$ and the identity theorem for complex numbers, we have

$$r_e(x; F_X^*) - \lambda x^2 = C_s - \lambda P, \quad \forall x \in \mathbb{R} \tag{51}$$

$$E[X^2] = P \tag{52}$$

We can show by substitution that (51)-(52) are satisfied when $\lambda = \frac{\log(e)}{2}\left(\frac{1}{\sigma_B^2+P} - \frac{1}{\sigma_E^2+P}\right)$ and F_X is selected as the Gaussian distribution with zero-mean and variance P, i.e., $F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi P}} e^{-\frac{y^2}{2P}} dy$. In this case, $C_s = \frac{1}{2}\log\left(1+\frac{P}{\sigma_B^2}\right) - \frac{1}{2}\log\left(1+\frac{P}{\sigma_E^2}\right)$. We claim that (51)-(52) cannot have another solution. To prove this claim, we note that (51)-(52) are independent of the amplitude constraint A and therefore are valid for any A, in particular for $A \to \infty$. That is, (51)-(52) are also the KKT conditions for the amplitude unconstrained problem (c.f. Appendix B):

$$\max_{E[X^2] \leq P} I(X; Y) - I(X; Z) \tag{53}$$

It is well-known by [3] using the entropy power inequality or alternatively by [17] using the I-MMSE relation that the unique solution of (53) is the Gaussian input distribution with zero-mean and variance P. We conclude that whenever (45)-(47) have a solution F_X^* with a support set of infinitely many points, it is a Gaussian distribution. However, since Gaussian distribution does not satisfy the amplitude constraint, the optimal input distribution F_X^* that achieves the secrecy capacity C_s cannot have infinitely many mass points, and must be a discrete distribution with finite support.

We can extend this contradiction argument for the entire rate-equivocation region. Consider the optimization problem for determining the boundary point of the rate-equivocation region with slope $\mu \geq 0$:

$$\max_{F_X \in \Omega_{A,P}} (\mu + 1) I(X; Y) - I(X; Z) \tag{54}$$

Note that if the variance constraint is not tight, i.e., $E[X^2] < P$, the problem again reduces to the case where only the amplitude constraint is present, in which case the optimal input distribution is discrete with finite support by Theorem 2. Hence, we assume without loss of generality that the variance constraint is tight and the necessary and sufficient optimality conditions for (54) are:

$$\mu i_B(x; F_X^*) + r_e(x; F_X^*) - \lambda x^2 \leq (\mu + 1) I_B(F_X^*) - I_E(F_X^*) - \lambda P, \quad \forall x \in [-A, A] \tag{55}$$

$$\mu i_B(x; F_X^*) + r_e(x; F_X^*) - \lambda x^2 \leq (\mu + 1) I_B(F_X^*) - I_E(F_X^*) - \lambda P, \quad \forall x \in \mathcal{S}_{F_X^*} \tag{56}$$

$$E[X^2] = P \tag{57}$$

Next, we prove by contradiction that the input distribution F_X^* that satisfies (55)-(57) must be a discrete distribution with finite support. Assume that $\mathcal{S}_{F_X^*}$ has infinite number of elements. In view of (55)-(57), analyticity of $i_B(z; F_X)$, $r_e(z; F_X)$ and z^2 over $\mathbb{C}$ and the identity theorem for complex numbers, we have

$$\mu i_B(x; F_X^*) + r_e(x; F_X^*) - \lambda x^2 = (\mu + 1) I_B(F_X^*) - I_E(F_X^*) - \lambda P, \quad \forall x \in \mathbb{R} \tag{58}$$

$$E[X^2] = P \tag{59}$$

It is easy to verify by substitution that (58)-(59) are satisfied when $\lambda = \frac{\log(e)}{2}\left(\frac{\mu+1}{\sigma_B^2+P} - \frac{1}{\sigma_E^2+P}\right)$ and F_X is selected as the Gaussian distribution with zero-mean and variance P. In this case, $I_B(F_X^*) = \frac{1}{2}\log\left(1+\frac{P}{\sigma_B^2}\right)$ and $I_E(F_X^*) = \frac{1}{2}\log\left(1+\frac{P}{\sigma_E^2}\right)$. Moreover, as in the secrecy capacity case, (58)-(59) cannot have another solution since (58)-(59) are independent of the amplitude constraint A and therefore are valid for $A \to \infty$. That is, (58)-(59) are also the KKT conditions for the amplitude unconstrained problem

$$\max_{E[X^2] \leq P} (\mu + 1) I(X; Y) - I(X; Z) \tag{60}$$

It is known from [3] and [17] that for all $\mu \geq 0$ the unique solution of (60) is also the Gaussian input distribution with zero-mean and variance P. This causes a contradiction since Gaussian input distribution is not amplitude constrained. Therefore, F_X^* is discrete with finite support. The two parts in this section prove the following theorem.

Theorem 4: Let F_X^* be the distribution that attains the secrecy capacity of the Gaussian wiretap channel with peak and average power constraints. The support set $\mathcal{S}_{F_X^*}$ is a finite set. More generally, the support set of distributions that attain the boundary of the entire rate-equivocation region under amplitude and variance constraints are finite sets.

We now provide an illustration for the effect of the variance constraint on the secrecy capacity achieving input distribution. Let $\sigma_B^2 = 1$, $\sigma_E^2 = 2$ and $A = 1$. If the variance constraint is $P \geq 1$, it is inactive for any input distribution, i.e., the problem reduces to the one with amplitude constraint only. In this case, in view of Theorem 3, the optimal distribution is the symmetric binary distribution $\frac{1}{2}\delta_{-A} + \frac{1}{2}\delta_A$. We now impose a variance

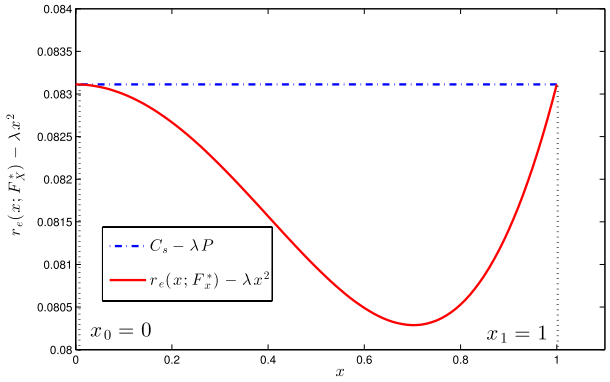

Fig. 6. The KKT conditions for $\sigma_B^2=1$, $\sigma_E^2=2$, $A=1$ and $P=0.8$. Equality is satisfied at points $\pm A$ and 0. Note that the plot is symmetric around $x=0$ and is shown only for $x\geq 0$.

constraint $P=0.8$. Clearly, in this case, the symmetric binary distribution at $\pm A$ is not feasible. We numerically find that the symmetric ternary distribution $0.4\delta_{-A}+0.2\delta_0+0.4\delta_A$ is optimal in this case and the corresponding Lagrange multiplier is $\lambda=0.116753$. We provide the plot of the KKT condition in Fig. 6 where we observe that $r_e(x;F_X^*)-\lambda x^2$ always lies below $C_s-\lambda P$ with equality on the support set.

VI. Conclusion

In this paper, we study the Gaussian wiretap channel under peak and average power constraints. We show that the boundary of the entire rate-equivocation region is achieved by discrete input distributions that have finite support. We prove this result by using the methodology in [4] and [5] for our setting. An interesting aspect that our result reveals is that, unlike the average power constrained Gaussian wiretap channel, under a peak power constraint, the secrecy capacity and the capacity cannot be obtained simultaneously in general, i.e., there is a tradeoff between the rate and the equivocation for the peak power constrained case. In the special case $A\leq 1.05$, we show that the secrecy capacity and the capacity are achieved simultaneously by a symmetric binary distribution at $\pm A$. Finally, we extend the discreteness result for the case when we have both amplitude and variance constraints.

Appendix A
Proof of Lemma 2

We first note that $p_Z(y)>\phi_E(|y|+A)$. We divide the integral into the following two regions $(-\infty,y]$, (y,∞) and apply this bound to obtain

$$\begin{aligned}&\int_{\mathbb{R}}\phi_D(\tau)\log\left(p_Z(y-\tau)\right)d\tau\\&\geq-\int_{-\infty}^{y}\phi_D(\tau)\frac{\log(e)(y-\tau+A)^2}{2\sigma_E^2}d\tau\\&\quad-\int_{y}^{\infty}\phi_D(\tau)\frac{\log(e)(\tau-y+A)^2}{2\sigma_E^2}d\tau+\log\left(\frac{1}{\sqrt{2\pi\sigma_E^2}}\right)\end{aligned}\tag{61}$$

Rearranging this bound, we get

$$\begin{aligned}&\int_{\mathbb{R}}\phi_D(\tau)\log\left(p_Z(y-\tau)\right)d\tau\\&\geq\log\left(\frac{1}{\sqrt{2\pi\sigma_E^2}}\right)-\log(e)\int_{-\infty}^{\infty}\phi_D(\tau)\frac{(y-\tau)^2+A^2}{2\sigma_E^2}d\tau\\&\quad-\frac{A}{\sigma_E^2}\log(e)\left(\int_{y}^{\infty}(\tau-y)\phi_D(\tau)d\tau+\int_{-\infty}^{y}(y-\tau)\phi_D(\tau)d\tau\right)\end{aligned}\tag{62}$$

$$\begin{aligned}&=\log\left(\frac{1}{\sqrt{2\pi\sigma_E^2}}\right)-\frac{\log(e)}{2\sigma_E^2}\left(y^2+A^2+\sigma_E^2-\sigma_B^2\right)\\&\quad-\frac{A\log(e)}{\sigma_E^2}b(y)\end{aligned}\tag{63}$$

where

$$b(y)=\int_{y}^{\infty}(\tau-y)\phi_D(\tau)d\tau+\int_{-\infty}^{y}(y-\tau)\phi_D(\tau)d\tau\tag{64}$$

$$=y\left(1-2Q\left(\frac{y}{\sigma_D}\right)\right)+\frac{2\sigma_D}{\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_D^2}}\tag{65}$$

where $\sigma_D^2=\sigma_E^2-\sigma_B^2$ and $Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$. We note that $b(y)\in o(y^2)$, i.e., $\frac{b(y)}{y^2}\to 0$ as $y\to\infty$ due to the fact that $Q(x)\leq 1$ and $e^{-y^2}\leq 1$.

On the other hand, we have $p_Y(y)\leq\phi_B(y-A)$ for $y>A$. Therefore,

$$\log\left(p_Y(y)\right)\leq\log\left(\frac{1}{\sqrt{2\pi\sigma_B^2}}\right)-\frac{(y-A)^2}{2\sigma_B^2}\log(e),\quad y>A\tag{66}$$

Consequently, in order to prove the asserted inequality in (30), it suffices to show that there exists y' sufficiently large such that for all $y>y'$

$$\begin{aligned}&\log\left(\frac{1}{\sqrt{2\pi\sigma_B^2}}\right)-\frac{(y-A)^2}{2\sigma_B^2}\log(e)\\&\leq\log\left(\frac{1}{\sqrt{2\pi\sigma_E^2}}\right)-\frac{\left(y^2+A^2+\sigma_E^2-\sigma_B^2\right)}{2\sigma_E^2}\log(e)\\&\quad-\frac{A\log(e)}{\sigma_E^2}b(y)\end{aligned}\tag{67}$$

As $b(y)\in o(y^2)$, (67) is equivalent to

$$-\frac{y^2}{\sigma_B^2}\leq-\frac{y^2}{\sigma_E^2}+o(y^2)\tag{68}$$

Since $\sigma_B^2<\sigma_E^2$, (68), and hence (67), is true for $y>y'$ for sufficiently large y'. This completes the proof of Lemma 2.

Appendix B
A Modified Proof for the AWGN Channel

In this section, we present a modified version of the discreteness proof in [4] and [5] for the AWGN channel under

amplitude and variance constraints. Our proof method closely follows the one in [4] and [5], but it takes a short-cut by directly relating the amplitude and variance constrained problem to the amplitude unconstrained but variance constrained problem. This is more readily generalizable to the Gaussian wiretap channel as done in Section V.

Consider the AWGN channel:

$$Y = X + N \tag{69}$$

where N is Gaussian with zero-mean and unit-variance. The channel capacity C of the AWGN channel under amplitude constraint A and variance constraint P is

$$C = \max_{F_X \in \Omega_{A,P}} I(X;Y) \tag{70}$$

where the feasible set of input distributions $\Omega_{A,P}$ is:

$$\Omega_{A,P} = \left\{ F_X : \int_{-A}^{A} dF_X(x) = 1, \quad \int_{-A}^{A} x^2 dF_X(x) \le P \right\} \tag{71}$$

By the Lagrangian theorem, $F_X^* \in \Omega_{A,P}$ is optimal if and only if there exists $\lambda \ge 0$ such that F_X^* is the unique solution of the following optimization problem:

$$\max_{F_X \in \Omega_A} I(X;Y) - \lambda E[X^2] \tag{72}$$

where

$$\Omega_A = \left\{ F_X : \int_{-A}^{A} dF_X(x) = 1 \right\} \tag{73}$$

Since the objective function in (72) is strictly concave in the input distribution, the directional derivative of the objective function with respect to F_X gives us the following necessary and sufficient conditions that the optimal input distribution F_X^* should satisfy [4], [5], [11]

$$i(x;F_X^*) - \lambda x^2 \le C - \lambda E[X^2], \quad \forall x \in [-A, A] \tag{74}$$

$$i(x;F_X^*) - \lambda x^2 = C - \lambda E[X^2], \quad \forall x \in \mathcal{S}_{F_X^*} \tag{75}$$

$$\lambda \left(E[X^2] - P \right) = 0 \tag{76}$$

We will show the discreteness of the optimal input distribution satisfying the KKT conditions in (74)-(76) by contradiction. To this end, we first note that when the second moment constraint in (76) is not active, i.e., when $\lambda = 0$, the problem reduces to the AWGN channel with only an amplitude constraint, for which we know that the optimal input distribution is discrete. Hence, from now on, we will focus on the case where the second moment constraint in (76) is active, i.e., $E\left[X^2\right] = P$. When this equality is satisfied, we can rewrite the KKT conditions in (74)-(76) as follows:

$$i(x;F_X^*) - \lambda x^2 \le C - \lambda P, \quad \forall x \in [-A, A] \tag{77}$$

$$i(x;F_X^*) - \lambda x^2 = C - \lambda P, \quad \forall x \in \mathcal{S}_{F_X^*} \tag{78}$$

$$E[X^2] = P \tag{79}$$

Now, we prove that the optimal input distribution satisfying (77)-(79) should be discrete by contradiction. To this end, we assume that the support set $\mathcal{S}_{F_X^*}$ includes infinitely many points. In view of the analyticity of $i(z;F_X^*)$ and z^2 over all complex numbers $\mathbb{C}$, this assumption implies that we should have

$$i(x;F_X^*) - \lambda x^2 = C - \lambda P, \quad \forall x \in \mathbb{R} \tag{80}$$

$$E[X^2] = P \tag{81}$$

for the optimal input distribution. We can verify by substitution that (80)-(81) are satisfied when $\lambda = \frac{\log(e)}{2(1+P)}$ and F_X is selected as the cumulative distribution function corresponding to the Gaussian density with zero-mean and variance P, i.e., $F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi P}} e^{-\frac{y^2}{2P}} dy$. In this case, $C = \frac{1}{2}\log(1+P)$.

Next, we claim that (80)-(81) cannot have another solution. To prove this claim, we note that (80)-(81) is independent of the amplitude constraint A and therefore is valid for any A and in particular $A \to \infty$. That is, (80)-(81) are also the KKT conditions for the amplitude unconstrained problem, i.e.,

$$\max_{E[X^2] \le P} I(X;Y) \tag{82}$$

It is well known (see [19]) that the unique optimal input distribution for (82) is Gaussian with zero-mean and variance P and the optimal value for (82) is $\frac{1}{2}\log(1+P)$. Consequently, (80)-(81) have a unique solution, which is $\lambda = \frac{\log(e)}{2(1+P)}$ and F_X is Gaussian with zero-mean and variance P. However, this causes a contradiction in view of the assumption that the input is amplitude constrained. Therefore, F_X^* is a discrete distribution with finite support.

REFERENCES

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

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

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

[4] J. G. Smith, "The information capacity of amplitude- and variance-constrained sclar Gaussian channels," *Inf. Control*, vol. 18, no. 3, pp. 203–219, Apr. 1971.

[5] J. G. Smith, "On the information capacity of peak and average power constrained Gaussian channels," Ph.D. dissertation, Dept. Elect. Eng., Univ. California, Berkeley, Berkeley, CA, USA, 1969.

[6] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1060–1071, Jul. 1995.

[7] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[8] A. Tchamkerten, "On the discreteness of capacity-achieving distributions," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2773–2778, Nov. 2004.

[9] M. C. Gursoy, H. V. Poor, and S. Verdu, "The noncoherent Rician fading channel—Part I: Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2193–2206, Sep. 2005.

[10] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2336–2351, Jul. 2005.

[11] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, Jun. 2005.

[12] M. Raginsky, "On the information capacity of Gaussian channels under small peak power constraints," in *Proc. Allerton Conf.*, Sep. 2008, pp. 286–293.

[13] L. Zhang and D. Guo, "Capacity of Gaussian channels with duty cycle and power constraints," in *Proc. IEEE ISIT*, Jul./Aug. 2011, pp. 513–517.
[14] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 712–714, Mar. 1997.
[15] N. Sharma and S. Shamai, "Transition points in the capacity-achieving distribution for the peak-power limited AWGN and free-space optical intensity channels," *Problems Inf. Trans.*, vol. 46, no. 4, pp. 283–299, Dec. 2010.
[16] D. Guo, S. Verdu, and S. Shamai, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
[17] D. Guo, Y. Wu, S. Verdu, and S. Shamai, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.
[18] G. Casella and W. E. Strawderman, "Estimating a bounded normal mean," *Ann. Statist.*, vol. 9, no. 4, pp. 870–878, Jul. 1981.
[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2006.

Omur Ozel (S'08–M'15) received B.Sc. and M.S. degrees with honors in Electrical and Electronics Engineering from the Middle East Technical University (METU), Ankara, Turkey in 2007 and 2009, respectively. He received his Ph.D. degree in Electrical and Computer Engineering (ECE) from the University of Maryland (UMD), College Park in December 2014. Since January 2015, he is a postdoctoral scholar in the Electrical Engineering and Computer Sciences Department at the University of California Berkeley. His doctoral research was awarded a Distinguished Dissertation Fellowship by the ECE Department and the second place award in the Clark School of Engineering Annual Doctoral Research Competition at UMD. His research interests lie in the intersection of wireless communications, information theory and networking.

Ersen Ekrem received his Ph.D. degree from the Department of Electrical and Computer Engineering at the University of Maryland, College Park in August 2012. Prior to that, he received the B.S. and M.S. degrees in Electrical and Electronics Engineering from Bogazici University, Istanbul, Turkey, in 2006 and 2007, respectively. Currently, he is with Qualcomm, Santa Clara, USA.

He received the Distinguished Dissertation Fellowship from the ECE Department at the University of Maryland, College Park, in 2012. His research interests include information theory and wireless communications.

Sennur Ulukus (S'90–M'98) is a Professor of Electrical and Computer Engineering at the University of Maryland at College Park, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member at AT&T Labs-Research. She received her Ph.D. degree in Electrical and Computer Engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, and B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University. Her research interests are in wireless communication theory and networking, network information theory for wireless communications, signal processing for wireless communications, information theoretic physical layer security, and energy harvesting communications.

Dr. Ulukus received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, an 2005 NSF CAREER Award, the 2010-2011 ISR Outstanding Systems Engineering Faculty Award, and the 2012 George Corcoran Education Award. She served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY (2007-2010) and IEEE TRANSACTIONS ON COMMUNICATIONS (2003-2007). She served as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on wireless communications powered by energy harvesting and wireless energy transfer (2015), IEEE JOURNAL OF COMMUNICATIONS AND NETWORKS for the special issue on energy harvesting in wireless networks (2012), IEEE TRANSACTIONS ON INFORMATION THEORY for the special issue on interference networks (2011), IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on multiuser detection for advanced communication systems and networks (2008). She served as the TPC co-chair of the 2014 IEEE PIMRC, Communication Theory Symposium at 2014 IEEE Globecom, Communication Theory Symposium at 2013 IEEE ICC, Physical-Layer Security Workshop at 2011 IEEE Globecom, Physical-Layer Security Workshop at 2011 IEEE ICC, 2011 Communication Theory Workshop (IEEE CTW), Wireless Communications Symposium at 2010 IEEE ICC, Medium Access Control Track at 2008 IEEE WCNC, and Communication Theory Symposium at 2007 IEEE Globecom. She was the Secretary of the IEEE Communication Theory Technical Committee (CTTC) in 2007-2009.