Optical Wiretap Channel With Input-Dependent Gaussian Noise Under Peak- and Average-Intensity Constraints

Morteza Soltani, *Student Member, IEEE*, and Zouheir Rezki, *Senior Member, IEEE*

***Abstract*— This paper studies the optical wiretap channel with input-dependent Gaussian noise, in which the main distortion is caused by an additive Gaussian noise whose variance depends on the current signal strength. Subject to nonnegativity and peak-intensity constraints on the channel input, we first present a practical optical wireless communication scenario for which the considered wiretap channel is stochastically degraded. We then study the secrecy-capacity-achieving input distribution of this wiretap channel and prove it to be discrete with a finite number of mass points, one of them located at the origin. Moreover, we show that the entire rate-equivocation region of this wiretap channel is also obtained by discrete input distributions with a finite support. Similar to the case of the Gaussian wiretap channel under a peak-power constraint, here too, we observe that under nonnegativity and peak-intensity constraints, there is a tradeoff between the secrecy capacity and the capacity in the sense that both may not be achieved simultaneously. Furthermore, we prove the optimality of discrete input distributions in the presence of an additional average intensity constraint. Finally, we shed light on the asymptotic behavior of the secrecy capacity in the low- and high-intensity regimes. In the low-intensity regime, the secrecy capacity scales quadratically with the peak-intensity constraint. On the other hand, in the high-intensity regime, the secrecy capacity does not scale with the constraint.**

***Index Terms*— Optical wiretap channel, intensity modulation and direct detection, peak- and average-intensity constraints, low-intensity regime, high-intensity regime, input-dependent Gaussian noise.**

I. Introduction

OPTICAL wireless communication (OWC) is a promising technique for supporting high data-rate communication as a complementary or a backup technology to radio-frequency (RF) communications. It has numerous advantages in comparison to RF, including higher data-rates, more abundant unlicensed spectrum and being less demanding in terms of system infrastructure.

One of the most popular communication techniques used in OWC is the intensity modulation and direct detection (IM-DD) technique for its simplicity [1]. In this setup, the channel input modulates the intensity of the emitted light. Thus, the input signal is proportional to the light intensity and is nonnegative. The receiver is usually equipped with a photodetector (PD) which measures the intensity of the received light and generates a signal proportional to the detected intensity, corrupted by noise. The simplest existing channel model for OWC is the *free space* optical (FSO) channel, where the corrupting noise at the receiver is independent of the input signal [2]. To reflect a more accurate channel model for OWC, [3] assumes the corrupting noise to be dependent on the input signal (due to the random nature of photon emission in the laser diode) and derives asymptotic upper and lower bounds on the capacity under nonnegativity, peak- and average-intensity constraints. The work in [4] also focuses on the optical intensity channels with input-dependent Gaussian noise and proves that under peak- and average-intensity constraints, discrete input distributions with finite supports are capacity-achieving. Additionally, the authors in [17] provide a general upper bound for the channel capacity of the input-dependent Gaussian noise optical channel based on the discrete input distributions.

Exchanging confidential information over a communication medium (wired, wireless or optical) in the presence of unauthorized eavesdroppers has been always a challenging problem for system designers. This problem has been conventionally addressed by cryptographic encryption [5] without considering the imperfections introduced by the communication channel. In this model, the usage of *secret keys* is the main approach for having secure communication. Wyner [6], on the other hand, proves the possibility of secure communications without relying on encryption by introducing the stochastically degraded wiretap channel model.

The wiretap channels are studied with respect to the rate-equivocation region, which is defined as the set of rate pairs for which the transmitter can communicate confidential messages reliably with a legitimate receiver while ensuring a certain secrecy level against an eavesdropper [7]. For the class of degraded wiretap channels, it is established in [6] that there exists a single-letter characterization for the rate-equivocation region. Leung-Yan-Cheong and Hellman [8] study the Gaussian wiretap channel under an average power constraint and obtain a single-letter expression for the entire rate-equivocation region. Particularly, they show that under an average power constraint, the Gaussian distribution is the optimal input distribution for attaining both the capacity and the secrecy capacity with no compromise between the communication rate and the equivocation rate at the eaves-

Manuscript received October 8, 2017; revised March 9, 2018; accepted June 14, 2018. Date of publication June 28, 2018; date of current version September 13, 2018. This work was supported by the King Abdullah University of Science and Technology, under a competitive research Grant OSR-2016-CRG5-2958-01. This paper was presented at the 2018 International Zurich Conference.

The authors are with the Department of Electrical and Computer Engineering, University of Idaho, Moscow, ID 83844 USA (e-mail: msoltani@uidaho.edu; zrezki@uidaho.edu).

Communicated by A. Khisti, 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.2018.2851225

dropper. On the other hand, under a peak-power constraint, the work in [9] proves that the entire rate-equivocation region of the Gaussian wiretap channel is achieved by discrete input distributions with finite supports. More specifically, the secrecy-capacity-achieving input distribution may not be identical to the capacity-achieving counterpart in general, resulting in a tradeoff between the rate and its equivocation.

This work considers an *optical* wiretap channel with input-dependent Gaussian noise which consists of a transmitter, a legitimate user and an eavesdropper. We assume that the output signals at both the legitimate user's and the eavesdropper's channels are corrupted by both input-dependent and input-independent Gaussian noises. In this setup, the objective is to have a secure communication with the legitimate user over an optical channel while keeping the eavesdropper ignorant of the transmitted message as much as possible. We study the optical wiretap channel with input-dependent Gaussian noise under nonnegativity, peak- and average-intensity constraints. We first present a practical OWC scenario based on IM-DD technique for which the optical wiretap channel with input-dependent Gaussian noise is stochastically degraded. We then use the results in [6] to conclude that there exists a single-letter expression for the entire rate-equivocation region. Next, we employ the functional optimization problem addressed in [10] and [4] to obtain necessary and sufficient conditions, also known as Karush-Kuhn-Tucker (KKT) conditions, for the optimal input distribution. Using KKT conditions, we prove by contradiction that the secrecy capacity as well as the entire rate-equivocation region of this wiretap channel are obtained by discrete input distributions with a finite number of mass points. Finally, we provide an asymptotic analysis for the secrecy capacity in the low- and high-intensity regime. More specifically, we observe that in the low-intensity regime, the secrecy capacity is achieved by a binary input distribution and it scales quadratically with the peak-intensity constraint. In the high-intensity regime, the secrecy capacity can be upper-bounded by a constant value implying that it does not scale with the constraints. Our numerical results demonstrate that likewise the case of the Gaussian wiretap channel under a peak-power constraint, here too, the secrecy capacity and the capacity are not achieved by the same distribution in general. This, in turn, implies that there is a tradeoff between the rate and its equivocation.

In the case of the optical wiretap channel with input-dependent Gaussian noise, due to the existence of input-dependent noise components, our technical proofs differ from those of [9]. More specifically, our analysis for showing the analyticity of the mutual information densities are more challenging. Additionally, our contradiction statements for proving the discreteness of the optimal input distribution are different. Besides, we prove that the secrecy-capacity-achieving input distribution has a mass point at the origin for the case of peak-intensity constraint as well as the case of peak- and average-intensity constraints.

II. System Model

We consider a practical OWC system where IM-DD is employed for optical communication. In this setup, the channel input modulates the emitted light intensity from Light Emitting Diode (LED) at the transmitter and PDs are used for receiving the optical signal at the legitimate user's and eavesdropper's receivers. We assume that there exist line-of-sight (LoS) paths between the optical transmitter and the receivers. In such a scenario, the received power of the LoS path dominates the received power of the reflected paths. Hence, the optical wireless channel between the transmitter and the legitimate user and between the transmitter and the eavesdropper become LoS channels [11, Ch. 2]. Figure 1 shows the geometry of an LoS OWC link with arbitrary receiver orientation. The LoS optical wireless channels between the transmitter and the legitimate user and between the transmitter and the eavesdropper are denoted by positive reals h and g, respectively, and are given by [11, Ch. 2]

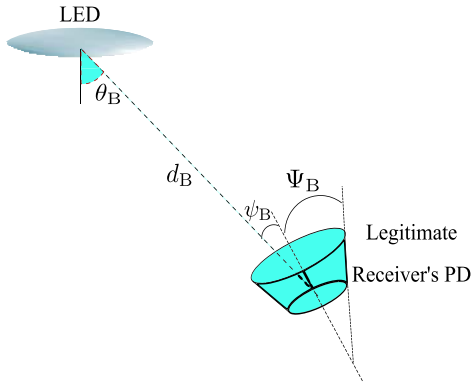

Fig. 1. Geometry of an LoS OWC link.

$$h = \begin{cases} \dfrac{(m+1)A_{\mathrm{B}}}{2\pi d_{\mathrm{B}}^2} \cos^m(\theta_{\mathrm{B}}) \cos(\psi_{\mathrm{B}})\, f_{\mathrm{B}}\, k_{\mathrm{B}}, & \text{if } 0 \le \psi_{\mathrm{B}} \le \Psi_{\mathrm{B}}, \\ 0, & \text{if } \psi_{\mathrm{B}} > \Psi_{\mathrm{B}}, \end{cases} \tag{1}$$

$$g = \begin{cases} \dfrac{(m+1)A_{\mathrm{E}}}{2\pi d_{\mathrm{E}}^2} \cos^m(\theta_{\mathrm{E}}) \cos(\psi_{\mathrm{E}})\, f_{\mathrm{E}}\, k_{\mathrm{E}}, & \text{if } 0 \le \psi_{\mathrm{E}} \le \Psi_{\mathrm{E}}, \\ 0, & \text{if } \psi_{\mathrm{E}} > \Psi_{\mathrm{E}}, \end{cases} \tag{2}$$

where m is the order of Lambertian emission, which depends on the semiangle at half illuminance of the LED $\Psi_{1/2}$ and is given by

$$m = -\frac{\ln(2)}{\ln\left(\cos\left(\Psi_{1/2}\right)\right)}. \tag{3}$$

The indices B and E stand for the legitimate user Bob and the eavesdropper Eve; d_{B} and d_{E} are the distances between the transmitter and the legitimate user and between the transmitter and the eavesdropper, respectively; A_{B} and A_{E} are the physical areas of the PDs at the legitimate user and the eavesdropper, respectively; $\theta_{\mathrm{B}} \in [0, \pi/2)$ and $\theta_{\mathrm{E}} \in [0, \pi/2)$ are the angles between the emitted light and the normal to the LED surface toward the legitimate receiver and the eavesdropper, respectively; ψ_{B} and ψ_{E} are the incident angle at the legitimate user and the eavesdropper, respectively; $\Psi_{\mathrm{B}} \in [0, \pi/2)$ and $\Psi_{\mathrm{E}} \in [0, \pi/2)$ are the concentrator field-of-view (FoV) of the

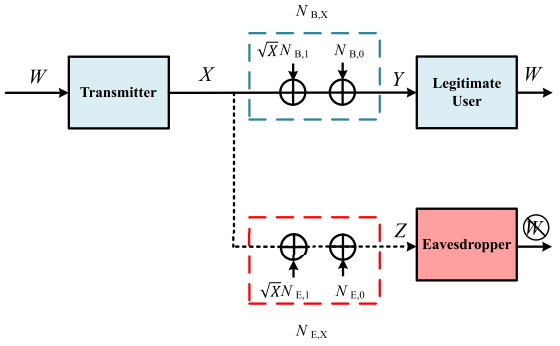

Fig. 2. The optical wiretap channel with input-dependent Gaussian noise.

legitimate user and the eavesdropper, respectively; f_{B}, k_{B}, f_{E} and k_{E} denote the optical filter gain and the concentrator gain of the legitimate user and the eavesdropper, respectively, and are assumed to be constant over their respective FoVs. An LoS IM-DD optical wireless link with input-dependent Gaussian noise from the transmitter to the legitimate user and from the transmitter to the eavesdropper can be respectively modeled as [11, Ch. 7]

$$\begin{cases} \widetilde{Y} = hX + \sqrt{hX}\widetilde{N}_{\mathrm{B},1} + \widetilde{N}_{\mathrm{B},1}, \\ \widetilde{Z} = gX + \sqrt{gX}\widetilde{N}_{\mathrm{E},1} + \widetilde{N}_{\mathrm{E},0}, \end{cases} \tag{4}$$

where X is the channel input ant it is a nonnegative random variable representing the intensity of the optical signal. Moreover, due to practical and safety restrictions, the input intensity is constrained by a peak constraint in general, i.e., $X \leq A$ [1]. Therefore, the channel input is constrained as

$$0 \leq X \leq A. \tag{5}$$

$\widetilde{Y}$ and $\widetilde{Z}$ are the received optical signals at the legitimate user and the eavesdropper, respectively. h and g are the LoS channel gains between the transmitter and the legitimate user and between the transmitter and the eavesdropper, respectively, and are given by (1)–(2). $\widetilde{N}_{\mathrm{B},0} \sim \mathcal{N}(0, \widetilde{\sigma}^2_{\mathrm{B},0})$ and $\widetilde{N}_{\mathrm{B},1} \sim \mathcal{N}(0, \widetilde{\sigma}^2_{\mathrm{B},1})$ are the input-independent and input-dependent Gaussian noise components at the legitimate receiver, respectively, and are assumed to be independent. Similarly, $\widetilde{N}_{\mathrm{E},0} \sim \mathcal{N}(0, \widetilde{\sigma}^2_{\mathrm{E},0})$ and $\widetilde{N}_{\mathrm{E},1} \sim \mathcal{N}(0, \widetilde{\sigma}^2_{\mathrm{E},1})$ are the input-independent and input-dependent Gaussian noise components at the eavesdropper, respectively and are assumed to be independent. Furthermore, $\widetilde{N}_{\mathrm{B},0}$ and $\widetilde{N}_{\mathrm{E},0}$ are also assumed to be independent. We note that the input-dependent noise in (4) is due to the nonlinearity in the optical channel [11, Ch. 7].

Figure 2 depicts an optical wiretap channel that is equivalent to the optical wiretap channel described by (4). In this equivalent wiretap channel, each link is an optical channel with input-dependent Gaussian noise model and is given by [3]

$$\begin{cases} Y = X + \sqrt{X}N_{\mathrm{B},1} + N_{\mathrm{B},0}, \\ Z = X + \sqrt{X}N_{\mathrm{E},1} + N_{\mathrm{E},0}, \end{cases} \tag{6}$$

where $Y = \frac{\widetilde{Y}}{h}$, $Z = \frac{\widetilde{Z}}{g}$, $N_{\mathrm{B},0} \sim \mathcal{N}(0, \sigma^2_{\mathrm{B}})$, $N_{\mathrm{B},1} \sim \mathcal{N}(0, \sigma^2_{\mathrm{B}}\eta^2_{\mathrm{B}})$, $N_{\mathrm{E},0} \sim \mathcal{N}(0, \sigma^2_{\mathrm{E}})$ and $N_{\mathrm{E},1} \sim \mathcal{N}(0, \sigma^2_{\mathrm{E}}\eta^2_{\mathrm{E}})$ with $\sigma^2_{\mathrm{B}} = \frac{\widetilde{\sigma}^2_{\mathrm{B},0}}{h^2}$, $\eta^2_{\mathrm{B}} = \frac{\widetilde{\sigma}^2_{\mathrm{B},1}}{\widetilde{\sigma}^2_{\mathrm{B},0}}h$, $\sigma^2_{\mathrm{E}} = \frac{\widetilde{\sigma}^2_{\mathrm{E},0}}{g^2}$, and $\eta^2_{\mathrm{E}} = \frac{\widetilde{\sigma}^2_{\mathrm{E},1}}{\widetilde{\sigma}^2_{\mathrm{E},0}}g$, respectively. The variables η^2_{B} and η^2_{E} denote the ratios of the input-dependent noise variances to the input-independent noise variances of the legitimate user's and the eavesdropper's channels, respectively. Notice that changing the orientation of the transmitter with respect to the legitimate receiver's and the eavesdropper's positions affects the channel gains h and g through ψ_{B}, θ_{B}, ψ_{E}, θ_{E}, respectively. We note that if the system parameters satisfy

$$\frac{\cos(\psi_{\mathrm{B}})}{\cos(\psi_{\mathrm{E}})}\left[\frac{\cos(\theta_{\mathrm{B}})}{\cos(\theta_{\mathrm{E}})}\right]^m = \frac{\widetilde{\sigma}^2_{\mathrm{B},1}}{\widetilde{\sigma}^2_{\mathrm{E},1}}\, d^2_{\mathrm{BE}}\, \rho_{\mathrm{EB}}, \tag{7}$$

where $d^2_{\mathrm{BE}} = \frac{d^2_{\mathrm{B}}}{d^2_{\mathrm{E}}}$ and $\rho_{\mathrm{EB}} = \frac{A_{\mathrm{E}} f_{\mathrm{E}} k_{\mathrm{E}}}{A_{\mathrm{B}} f_{\mathrm{B}} k_{\mathrm{B}}}$, then, using (1)–(2), we have

$$\sigma^2_{\mathrm{B}}\eta^2_{\mathrm{B}} = \sigma^2_{\mathrm{E}}\eta^2_{\mathrm{E}}, \tag{8}$$

which implies that the input-dependent noise components in both channels are statistically equivalent. One can show via numerical inspections that (8) might be satisfied for various system parameters θ_{B}, ψ_{B}, θ_{E}, ψ_{E}, $\widetilde{\sigma}^2_{\mathrm{B},1}$, $\widetilde{\sigma}^2_{\mathrm{E},1}$, d^2_{BE} and ρ_{EB}. This implies that in such cases, $N_{\mathrm{B},1}$ and $N_{\mathrm{E},1}$ are statistically equivalent and the optical wiretap channel with input-dependent Gaussian noise can be stochastically degraded.

In the sequel, without loss of generality, we focus on the wiretap channel described by (6) whose input is X and whose outputs are Y and Z, at the legitimate receiver and the eavesdropper, respectively. In this optical wiretap channel, if $\sigma^2_{\mathrm{B}} < \sigma^2_{\mathrm{E}}$ and $\sigma^2_{\mathrm{B}}\eta^2_{\mathrm{B}} = \sigma^2_{\mathrm{E}}\eta^2_{\mathrm{E}}$, then the random variables X, Y and Z form the Markov chain $X \to Y \to Z$ and consequently the optical wiretap channel becomes stochastically degraded. As a result, the rate-equivocation region of such an optical channel can be expressed in a single-letter form due to [6]. Furthermore, under the conditions $\sigma^2_{\mathrm{B}} \geq \sigma^2_{\mathrm{E}}$ and $\sigma^2_{\mathrm{B}}\eta^2_{\mathrm{B}} = \sigma^2_{\mathrm{E}}\eta^2_{\mathrm{E}}$, the random variables X, Y and Z form the Markov chain $X \to Z \to Y$, from which it can be easily inferred that the secrecy capacity (defined later in this section) is equal to zero.

The Rate-Equivocation Characterization of the Optical Wiretap Channel With Input-Dependent Gaussian Noise

An $(n, 2^{nR})$ code for the peak-intensity-constrained optical wiretap channel with input-dependent Gaussian noise consists of the random variable W (message set) uniformly distributed over the set $\mathcal{W} = \{1, 2, \ldots, 2^{nR}\}$, an encoder at the transmitter $f_n : \mathcal{W} \to [0, A]^n$ satisfying the nonnegativity and peak-intensity constraints, and a decoder at the legitimate user $g_n : \mathbb{R}^n \to \mathcal{W}$. The equivocation of the confidential message W is defined as the eavesdropper's uncertainty about W and is measured by the normalized conditional entropy $\frac{1}{n}H(W|Z^n)$. The probability of error for such a code is defined as $P^n_{\mathrm{e}} = \Pr\{g_n(Y^n) \neq W\}$. A rate-equivocation pair (R, R_{eq}) is said to be achievable if there exists an $(n, 2^{nR})$ code satisfying

$$\lim_{n\to\infty} P^n_{\mathrm{e}} = 0, \tag{9}$$

$$R_{\mathrm{eq}} \leq \lim_{n\to\infty} \frac{1}{n} H(W|Z^n). \tag{10}$$

The rate-equivocation region consists of all achievable rate-equivocation pairs, and is denoted by $\mathcal{E}$. A rate R is said to

be perfectly secure if we have $R_{\text{eq}} = R$, i.e., if there exists an $(n, 2^{nR})$ code satisfying $\lim_{n\to\infty} \frac{1}{n} I(W; Z^n) = 0$. The supremum of such rates is defined to be the secrecy capacity and denoted by C_{S}.

Since under the assumption of $\sigma_{\text{B}}^2\eta_{\text{B}}^2 = \sigma_{\text{E}}^2\eta_{\text{E}}^2$ and $\sigma_{\text{E}}^2 > \sigma_{\text{B}}^2$, the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints is stochastically degraded, its entire rate-equivocation region $\mathcal{E}$ can be expressed in a single-letter expression and the entire rate-equivocation region of the this wiretap channel is given by the union of the rate-equivocation pairs (R, R_{eq}) such that [6]

$$R \le I(X; Y), \tag{11}$$

$$R_{\text{eq}} \le I(X; Y) - I(X; Z), \tag{12}$$

for some input distribution $F_X \in \mathcal{A}^+$, where $I(X; Y)$ and $I(X; Z)$ are the mutual information of the legitimate user's and the eavesdropper's channels, respectively, and the feasible set $\mathcal{A}^+$ is given by

$$\mathcal{A}^+ \triangleq \left\{ F_X : \int_0^A dF_X(x) = 1 \right\}. \tag{13}$$

III. Main Results

In this section, we present the main results related to the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints. We first focus on the secrecy capacity and prove the discreteness of the secrecy-capacity-achieving input distribution. We then establish that the entire rate-equivocation region of this wiretap channel is also obtained by discrete input distributions with finite supports.

A. Results on the Secrecy Capacity

The secrecy capacity of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints is given by the solution of the following optimization problem

$$\max_{F_X \in \mathcal{A}^+} g_0(F_X), \tag{14}$$

where $g_0(F_X) = I(X; Y) - I(X; Z)$. Under the constraint (5), the solution of (14) is discrete with a finite support as stated by Theorem 1.

Theorem 1: There exists a unique input distribution that attains the secrecy capacity of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints. Furthermore, the support set of this optimal input distribution is a finite set.

Proof: The proof is provided in Section IV. ■

To prove Theorem 1, we first prove that the set of input distributions $\mathcal{A}^+$ that satisfies (13), is compact and convex. We then show that the objective function in (14) is continuous, strictly concave and weakly differentiable in the input distribution F_X and hence we conclude that the solution to the optimization problem (14) exists and is unique. We continue the proof by deriving the necessary and sufficient conditions (KKT conditions) for the optimality of the optimal input distribution F_X^* and finally by means of contradiction we show that this optimal input distribution is discrete with a finite number of mass points. Unlike the case of the Gaussian wiretap channel under a peak-power constraint, where the corrupting noise components are assumed to be independent from channel input [9], the existence of input-dependent noise components in the optical wiretap channel with input-dependent Gaussian noise results in several challenging problems: 1) The conditions under which this wiretap channel under a peak-intensity constraint becomes stochastically degraded is different than that of [9]; 2) The proof of the analyticity of the mutual information density functions are different from those presented in [9]; 3) The technical steps for proving the discreteness of the solution to problem (14) are different from those utilized in [9] and cannot be generalized to those cases.

Next, we establish the existence of a mass point at $x = 0$ in the support set of the secrecy-capacity-achieving input distribution.

Proposition 1: Let $\mathcal{S}_{F_X^*}$ be the support set of the secrecy-capacity-achieving input distribution F_X^* for the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints. Then $x = 0$ always belongs to $\mathcal{S}_{F_X^*}$.

Proof: The proof is presented in Section IV. ■

It is worth mentioning that the existence of a mass point at the origin has also been established in [4] for the optical channel with input-dependent Gaussian noise, but with no secrecy constraints. Furthermore, the proof of Proposition 1 also holds true for the case where $\eta_{\text{B}}^2 = \eta_{\text{E}}^2 = 0$, i.e., the optical wiretap channel with input-independent Gaussian noise.

B. Results on the Rate-Equivocation Region

By a time-sharing argument, it can be shown that the rate-equivocation region of the optical wiretap channel with input-dependent Gaussian noise is convex. Therefore, the region can be characterized by finding tangent lines to $\mathcal{E}$, which are given by the solutions of

$$\max_{F_X \in \mathcal{A}^+} g_\lambda(F_X), \tag{15}$$

where $g_\lambda(F_X) = \lambda I(X; Y) + (1 - \lambda)[I(X; Y) - I(X; Z)]$, for all $\lambda \in [0, 1]$. Next, we establish that the entire rate-equivocation region of the optical wiretap channel with input-dependent Gaussian noise under the constraints (5) is also obtained by discrete input distributions with finite supports.

Theorem 2: There exists a unique input distribution that achieves the boundary of the rate-equivocation region of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints. This optimal input distribution is discrete with a finite support.

Proof: Theorem 2 is established in Section IV. ■

It is worth noting that for the case when $\eta_{\text{B}}^2 = \eta_{\text{E}}^2 = 0$ (i.e., for the optical wiretap channel with input-independent Gaussian noise), an approach similar to the one in [9] can be used to prove the discreteness of the optimal solutions (14) and (15). An interesting observation is that our contradiction argument for proving the discreteness of the optimal solutions of (14) and (15) (equations (60)–(69)), when $\eta_{\text{B}}^2, \eta_{\text{E}}^2 \neq 0$

cannot be generalized to the case when $\eta_{\mathrm{B}}^2 = \eta_{\mathrm{E}}^2 = 0$. A similar observation has been made in [4], but for the case with no secrecy constraint.

IV. Proof of the Main Results

In this section, we first provide the required preliminaries for the development of the main results. We then give the detailed proofs of the theorems stated in Section III.

A. Preliminaries and Notation

Since both channels are AWGN with input-dependent noise, the output densities for Y and Z exist for any input distribution F_X, and are given by

$$p_Y(y; F_X) = \int_0^A p(y|x)\, dF_X(x), \quad y \in \mathbb{R}, \tag{16}$$

$$p_Z(z; F_X) = \int_0^A p(z|x)\, dF_X(x), \quad z \in \mathbb{R}, \tag{17}$$

where $p(y|x)$ and $p(z|x)$ are given by [3]

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}} \exp\left(-\frac{(y-x)^2}{2\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}\right), \tag{18}$$

$$p(z|x) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}} \exp\left(-\frac{(z-x)^2}{2\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}\right). \tag{19}$$

We define the rate-equivocation density $r_{\mathrm{eq}}(x; F_X)$ as

$$r_{\mathrm{eq}}(x; F_X) = i_{\mathrm{B}}(x; F_X) - i_{\mathrm{E}}(x; F_X), \tag{20}$$

where $i_{\mathrm{B}}(x; F_X)$ and $i_{\mathrm{E}}(x; F_X)$ are the mutual information densities for the legitimate user's and eavesdropper's channel, respectively, and are given by

$$i_{\mathrm{B}}(x; F_X) = -\int_{\mathbb{R}} p(y|x) \log\left(p_Y(y; F_X)\right)\, dy - \frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)), \tag{21}$$

$$i_{\mathrm{E}}(x; F_X) = -\int_{\mathbb{R}} p(z|x) \log\left(p_Z(z; F_X)\right)\, dz - \frac{1}{2}\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)). \tag{22}$$

The mutual information and the mutual information density are related through

$$I(X;Y) = \int_0^A i_{\mathrm{B}}(x; F_X)\, dF_X(x), \tag{23}$$

$$I(X;Z) = \int_0^A i_{\mathrm{E}}(x; F_X)\, dF_X(x). \tag{24}$$

Since the channel input X satisfies (5), it can be shown that the conditional densities in (18) and (19) can be bounded as [4, Lemma 3]

$$\exp(-\alpha - \beta' y^2) \le p(y|x) \le \exp(\alpha - \beta y^2), \tag{25}$$

$$\exp(-\mu - \xi' z^2) \le p(z|x) \le \exp(\mu - \xi z^2), \tag{26}$$

for all $x \in [0, A]$, $y, z \in \mathbb{R}$, where $\alpha, \beta, \beta', \mu, \xi$ and ξ' are positive constants. Hence, for all $F_X \in \mathcal{A}^+$

$$\exp(-\alpha - \beta' y^2) \le p_Y(y; F_X) \le \exp(\alpha - \beta y^2), \tag{27}$$

$$\exp(-\mu - \xi' z^2) \le p_Z(z; F_X) \le \exp(\mu - \xi z^2). \tag{28}$$

Thus, we can write

$$|\log\left(p_Y(y; F_X)\right)| \le \alpha + \beta' y^2, \tag{29}$$

$$|\log\left(p_Z(z; F_X)\right)| \le \mu + \xi' z^2. \tag{30}$$

Next, we prove Theorem 1 using the preliminaries provided in this section.

B. Proof of Theorem 1

1) The Feasible Set $\mathcal{A}^+$ Is Compact and Convex: The proof follows along similar lines as in [12, Appendix A.1].

2) $g_0(F_X)$ Is Continuous in F_X: In order to show that $g_0(F_X)$ is a continuous function in F_X, it is sufficient to show that $I(X;Y)$ is continuous in F_X. The continuity of $I(X;Z)$ in F_X can be shown by following along similar lines as those in the proof of the continuity of $I(X;Y)$ in F_X. To this end, let us consider a sequence $\{F_X^{(n)}\}_{n\in\mathbb{N}}$ in $\mathcal{A}^+$ such that $F_X^{(n)} \to F_X$ for some $F_X \in \mathcal{A}^+$. It is evident that $p(y|x)$ is a continuous and bounded function in x and y, thus,

$$\lim_{n\to\infty} p_Y(y; F_X^{(n)}) = \lim_{n\to\infty} \int_0^A p(y|x)\, dF_X^{(n)}(x) = \int_0^A p(y|x)\, dF_X(x), \tag{31}$$

where (31) follows by the Helly-Bray Theorem [13]. Then,

$$\lim_{n\to\infty} p(y|x) \log(p_Y(y; F_X^{(n)})) = p(y|x) \log\left(p_Y(y; F_X)\right). \tag{32}$$

Moreover, by observing (25) and (29), we conclude that

$$|p(y|x) \log(p_Y(y; F_X^{(n)}))| \le \exp(\alpha - \beta y^2)[\alpha + \beta' y^2]. \tag{33}$$

Since the right hand side of (33) is absolutely integrable, we have

$$\int_{-\infty}^{+\infty} |p(y|x) \log(p_Y(y; F_X^{(n)}))|\, dy < \infty. \tag{34}$$

Thus, by applying the Dominated Convergence Theorem, we get

$$\lim_{n\to\infty} -\int_{\mathbb{R}} p(y|x) \log(p_Y(y; F_X^{(n)}))\, dy = -\int_{\mathbb{R}} \lim_{n\to\infty} p(y|x) \log(p_Y(y; F_X^{(n)}))\, dy = -\int_{\mathbb{R}} p(y|x) \log(p_Y(y; F_X))\, dy. \tag{35}$$

Additionally, $\frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x))$ is a bounded and continuous function for all $x \in [0, A]$. Therefore, we conclude that $i_{\mathrm{B}}(x; F_X)$ is a bounded and continuous function in F_X. Finally, applying the Helly-Bray Theorem results in

$$\lim_{n\to\infty} \int_0^A i_{\mathrm{B}}(x; F_X)\, dF_X^{(n)}(x) = \int_0^A i_{\mathrm{B}}(x; F_X)\, dF_X(x), \tag{36}$$

which implies that $I(X;Y)$ is continuous in F_X. Similar steps lead to the fact that $I(X;Z)$ is also a continuous function in F_X. This further implies that the objective function $g_0(F_X)$ is continuous in F_X.

3) $g_0(F_X)$ Is Strictly Concave in F_X: To show that $g_0(F_X)$ is a strictly concave function in F_X, we first note that $g_0(F_X) = I(X;Y|Z)$ when random variables X, Y and Z form the Markov chain $X \to Y \to Z$. Next, we present a lemma that establishes that $I(X;Y|Z)$ is a strictly concave function in F_X.

Lemma 1: If the random variables X, Y and Z form the Markov chain $X \to Y \to Z$, then the conditional mutual information $I(X;Y|Z)$ is a strictly concave function in input distribution F_X. Furthermore, the output distributions are unique, i.e., if F_{X_1} and F_{X_2} are both secrecy-capacity-achieving, then $p_Y(y;F_{X_1}) = p_Y(y;F_{X_2})$ and $p_Z(z;F_{X_1}) = p_Z(z;F_{X_2})$.

Proof: See Appendix A ∎

4) $g_0(F_X)$ Is Weakly Differentiable: Defining $F_{X_\theta} = (1-\theta)\,F_{X_0} + \theta F_X,\ \forall F_X \in \mathcal{A}^+,\ \theta \in [0,1]$, we have to show that the following limit exists

$$\lim_{\theta\to 0} \frac{g_0((1-\theta)\,F_{X_0} + \theta F_X) - g_0(F_{X_0})}{\theta}. \tag{37}$$

Substituting (21) and (22) into (37), we get

$$\begin{aligned}
&\lim_{\theta\to 0}\Bigg[\frac{\int_0^A i_{\mathrm{B}}(x;F_{X_\theta})\,dF_{X_\theta} - \int_0^A i_{\mathrm{E}}(x;F_{X_\theta})\,dF_{X_\theta}}{\theta} \\
&\quad + \frac{\int_0^A i_{\mathrm{E}}(x;F_{X_0})\,dF_{X_0} - \int_0^A i_{\mathrm{B}}(x;F_{X_0})\,dF_{X_0}}{\theta}\Bigg] \\
&= \lim_{\theta\to 0}\Bigg[\frac{(1-\theta)\int_0^A i_{\mathrm{B}}(x;F_{X_\theta})\,dF_{X_0} + \theta\int_0^A i_{\mathrm{B}}(x;F_{X_\theta})\,dF_X}{\theta} \\
&\quad - \frac{(1-\theta)\int_0^A i_{\mathrm{E}}(x;F_{X_\theta})\,dF_{X_0} + \theta\int_0^A i_{\mathrm{E}}(x;F_{X_\theta})\,dF_X}{\theta} \\
&\quad - \frac{\int_0^A i_{\mathrm{B}}(x;F_{X_0})\,dF_{X_0} - \int_0^A i_{\mathrm{E}}(x;F_{X_0})\,dF_{X_0}}{\theta}\Bigg] \\
&= \lim_{\theta\to 0}\Bigg[\frac{\int_0^A \left[i_{\mathrm{B}}(x;F_{X_\theta}) - i_{\mathrm{B}}(x;F_{X_0})\right]dF_{X_0}}{\theta} \\
&\quad + \int_0^A i_{\mathrm{B}}(x;F_{X_\theta})\,dF_X - \int_0^A i_{\mathrm{B}}(x;F_{X_\theta})\,dF_{X_0} \\
&\quad - \frac{\int_0^A \left[i_{\mathrm{E}}(x;F_{X_\theta}) - i_{\mathrm{E}}(x;F_{X_0})\right]dF_{X_0}}{\theta} \\
&\quad - \int_0^A i_{\mathrm{E}}(x;F_{X_\theta})\,dF_X + \int_0^A i_{\mathrm{E}}(x;F_{X_\theta})\,dF_{X_0}\Bigg].
\end{aligned} \tag{38}$$

Next, we show that

$$\lim_{\theta\to 0} \frac{\int_0^A \left[i_{\mathrm{B}}(x;F_{X_\theta}) - i_{\mathrm{B}}(x;F_{X_0})\right]dF_{X_0}}{\theta} = 0, \tag{39}$$

$$\lim_{\theta\to 0} \frac{\int_0^A \left[i_{\mathrm{E}}(x;F_{X_\theta}) - i_{\mathrm{E}}(x;F_{X_0})\right]dF_{X_0}}{\theta} = 0. \tag{40}$$

To this end, we first prove (39). The proof of (40) follows along a similar line as that of (39). We start the proof by substituting $i_{\mathrm{B}}(x;F_X) = \int_{\mathbb{R}} p(y|x)\log\left(\frac{p(y|x)}{p_Y(y;F_X)}\right)dy$ into the left hand side of (39) to obtain

$$\lim_{\theta\to 0} \frac{-\int_0^A \int_{\mathbb{R}} p(y|x)\log\left(\dfrac{p_Y(y;F_{X_\theta})}{p_Y(y;F_{X_0})}\right)dy\,dF_{X_0}(x)}{\theta}. \tag{41}$$

By noting that $p_Y(y;F_{X_\theta}) = (1-\theta)\,p_Y(y;F_{X_0}) + \theta\,p_Y(y;F_X)$ and substituting this into (41), we get (42), as shown at the top of the next page, where (a) follows from the fact that when $\theta \to 0$, $\log(1+\theta) \to \theta$ and the limit exists. By substituting (39) and (40) into (38) and noting that $F_{X_\theta} \to F_{X_0}$ as $\theta \to 0$, (37) becomes

$$\begin{aligned}
&\int_0^A \left[i_{\mathrm{B}}(x;F_{X_0}) - i_{\mathrm{E}}(x;F_{X_0})\right]dF_X \\
&\quad - \int_0^A \left[i_{\mathrm{B}}(x;F_{X_0}) - i_{\mathrm{E}}(x;F_{X_0})\right]dF_{X_0} \\
&= \int_0^A r_{\mathrm{eq}}(x;F_{X_0})\,dF_X - g_0(F_{X_0}),
\end{aligned} \tag{43}$$

which implies that the objective function $g_0(F_X)$ is weakly differentiable. Since the feasible set $\mathcal{A}^+$ is compact and convex and the objective function $g_0(F_X)$ is continuous, strictly concave and weakly differentiable, steps analogous to [10, Corollary 1] [4, Th. 2] yield the following necessary and sufficient conditions for the optimality of the distribution F_X^*

$$r_{\mathrm{eq}}(x;F_X^*) \le C_{\mathrm{S}}, \quad \forall x \in [0,A], \tag{44}$$

$$r_{\mathrm{eq}}(x;F_X^*) = C_{\mathrm{S}}, \quad \forall x \in \mathcal{S}_{F_X^*}, \tag{45}$$

where $\mathcal{S}_{F_X^*}$ is the support set of F_X^* and the secrecy capacity C_{S} is expressed as

$$\begin{aligned}
C_{\mathrm{S}} = I_{\mathrm{B}}(F_X^*) - I_{\mathrm{E}}(F_X^*) &= h_Y(F_X^*) - h_Z(F_X^*) \\
&\quad + \frac{1}{2}\mathbb{E}_{F_X^*}\left[\log\left(\frac{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}\right)\right],
\end{aligned} \tag{46}$$

where $I_{\mathrm{B}}(F_X^*)$ and $I_{\mathrm{E}}(F_X^*)$ are the mutual information for Bob and Eve, respectively, generated by the optimal 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^*. Moreover, $\mathbb{E}_{F_X^*}$ denotes the expectation operator with respect to the optimal distribution F_X^*. We now prove by contradiction that the secrecy-capacity-achieving input distribution F_X^* has a finite number of mass points. To reach a contradiction, we use the KKT conditions in (44) and (45). To this end, we first show that both $i_{\mathrm{B}}(x;F_X)$ and $i_{\mathrm{E}}(x;F_X)$ have analytic extensions over some open connected set in the complex plane $\mathbb{C}$ that includes the nonnegative real line $\mathbb{R}_0^+$.

5) The Rate-Equivocation Density $r_{\mathrm{eq}}(x;F_X)$ Has an Analytic Extension to Some Open Connected Set in the Complex Plane $\mathbb{C}$: To prove the analyticity of $r_{\mathrm{eq}}(x;F_X)$ over some open connected set in the complex plane $\mathbb{C}$, it is sufficient to prove that $i_{\mathrm{B}}(x;F_X)$ has an analytic extension to the open connected set. Invoking similar steps as those in the proof of the analyticity of $i_{\mathrm{B}}(x;F_X)$, one can show that $i_{\mathrm{E}}(x;F_X)$ has also an analytic extension to the open connected set. We start by denoting $i_{\mathrm{B}}(w;F_X)$ as the extension of mutual information density of the legitimate user's channel to the complex plane. Now, we have

$$\begin{aligned}
i_{\mathrm{B}}(w;F_X) = &-\int_{\mathbb{R}} p(y|w)\log\left(p_Y(y;F_X)\right)dy \\
&- \frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w)),
\end{aligned} \tag{47}$$

$$\begin{aligned}
&\lim_{\theta\to 0}\frac{-\int_0^A\int_{\mathbb{R}} p(y|x)\log\left(\frac{p_Y(y;\,F_{X_\theta})}{p_Y(y;\,F_{X_0})}\right)dy\,dF_{X_0}(x)}{\theta}\\
&\quad=\lim_{\theta\to 0}\frac{-\int_0^A\int_{\mathbb{R}} p(y|x)\log\left(1+\theta\left[\frac{p_Y(y;\,F_X)}{p_Y(y;\,F_{X_0})}-1\right]\right)dy\,dF_{X_0}(x)}{\theta}\\
&\quad\overset{(a)}{=}\lim_{\theta\to 0}\frac{-\int_0^A\int_{\mathbb{R}} p(y|x)\,\theta\left[\frac{p_Y(y;\,F_X)}{p_Y(y;\,F_{X_0})}-1\right]dy\,dF_{X_0}(x)}{\theta}\\
&\quad=-\int_{\mathbb{R}}\int_0^A\frac{p_Y(y;\,F_X)}{p_Y(y;\,F_{X_0})}\,p(y|x)\,dF_{X_0}(x)\,dy\\
&\qquad+\int_{\mathbb{R}}\int_0^A p_Y(y|x)\,dF_{X_0}(x)\,dy\\
&\quad=\int_{\mathbb{R}} p_Y(y;\,F_{X_0})\,dy-\int_{\mathbb{R}}\frac{p_Y(y;\,F_X)}{p_Y(y;\,F_{X_0})}p_Y(y;\,F_{X_0})\,dy=1-1=0,
\end{aligned}\tag{42}$$

where w is the complex variable. It is clear that $\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2\ w))$ is analytic over $\mathcal{D}_1\overset{\triangle}{=}\left\{w:\Re(w)>\frac{-1}{\eta_{\mathrm{B}}^2}\right\}$, where $\Re(\cdot)$ is the real part of a complex variable. Similarly, $\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2\ w))$ is analytic over $\mathcal{D}_2\overset{\triangle}{=}\left\{w:\Re(w)>\frac{-1}{\eta_{\mathrm{E}}^2}\right\}$. Since $\eta_{\mathrm{B}}^2\sigma_{\mathrm{B}}^2=\eta_{\mathrm{E}}^2\sigma_{\mathrm{E}}^2$ and $\sigma_{\mathrm{E}}^2>\sigma_{\mathrm{B}}^2$, we have $\frac{-1}{\eta_{\mathrm{B}}^2}>\frac{-1}{\eta_{\mathrm{E}}^2}$. Defining $\mathcal{D}$ as $\mathcal{D}\overset{\triangle}{=}\mathcal{D}_1$, one can have both of the logarithm functions to be analytic over $\mathcal{D}$. We note that $\mathcal{D}$ is an open connected set in the complex plane $\mathbb{C}$. Next, we show that the continuation of $-\int_{\mathbb{R}} p(y|w)\log\left(p_Y(y;\,F_X)\right)dy$ to the complex plane is continuous over $\mathcal{D}$. To this end, let $\{w_n\}_{n\in\mathbb{N}}$ be a sequence of complex numbers in $\mathcal{D}$ converging to $w\in\mathcal{D}$, where $w_n=a_n+jb_n$. Since w_n converges, there exist a positive real $\delta>0$ such that $|w_n|<\delta$ and for some $n>N$. This further implies that $|b_n|<\delta$ for $n>N$. Now, let $\sigma^2_{\mathrm{B},X,\,r}(w_n)$ and $\sigma^2_{\mathrm{B},X,\,i}(w_n)$ be the real and imaginary parts of $\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2\ w_n)$, respectively, i.e.,

$$\sigma^2_{\mathrm{B},X,\,r}(w_n)=\Re(\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)),\tag{48}$$

$$\sigma^2_{\mathrm{B},X,\,i}(w_n)=\Im(\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)),\tag{49}$$

where $\Im(\cdot)$ is the imaginary part of a complex variable. We have

$$\begin{aligned}
|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2&=(\sigma^2_{B,X,\,r}(w_n))^2+(\sigma^2_{\mathrm{B},X,\,i}(w_n))^2\\
&\geq(\sigma^2_{\mathrm{B},X,\,r}(w_n))^2.
\end{aligned}\tag{50}$$

Since $w_n\in\mathcal{D}$, we have $\Re(w_n)>-1/\eta_{\mathrm{B}}^2$ and as a result, $\sigma^2_{\mathrm{B},X,\,r}(w_n)$ is a positive real value. Now, we can write

$$\begin{aligned}
&|p(y|w_n)|\\
&=\left|\frac{1}{\sqrt{2\pi\,\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)}}\exp\left(-\frac{(y-w_n)^2}{2\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)}\right)\right|\\
&\leq\frac{1}{\sqrt{2\pi\,\sigma^2_{\mathrm{B},X,\,r}(w_n)}}\left|\exp\left(-\frac{(y-a_n-jb_n)^2}{2(\sigma^2_{\mathrm{B},X,\,r}(w_n)+j\sigma^2_{\mathrm{B},X,\,i}(w_n))}\right)\right|\\
&=\frac{1}{\sqrt{2\pi\,\sigma^2_{\mathrm{B},X,\,r}(w_n)}}\exp\left(-\frac{\sigma^2_{\mathrm{B},X,\,r}(w_n)\left[(y-a_n)^2-b_n^2\right]}{2|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2}\right.\\
&\quad\left.+\frac{2\,b_n\,\sigma^2_{\mathrm{B},X,\,i}(w_n)\ (y-a_n)}{2|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2}\right)\\
&=\frac{1}{\sqrt{2\pi\,\sigma^2_{\mathrm{B},X,r}(w_n)}}\exp\left(\frac{b_n^2}{2\sigma^2_{\mathrm{B},X,r}(w_n)}\right)\\
&\quad\times\exp\left(-\frac{\sigma^2_{\mathrm{B},X,r}(w_n)\,(y-c_n)^2}{2|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2}\right)\\
&\leq\frac{1}{\sqrt{2\pi\,\sigma^2_{\mathrm{B},X,\,r}(w_n)}}\exp\left(\frac{\delta^2}{2\sigma^2_{\mathrm{B},X,\,r}(w_n)}\right)\\
&\quad\times\exp\left(-\frac{\sigma^2_{\mathrm{B},X,\,r}(w_n)\,(y-c_n)^2}{2|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2}\right)\\
&\leq M(\delta)\exp\left(-\frac{(y-c_n)^2}{d_n^2}\right),
\end{aligned}\tag{51}$$

where $M(\delta)$ is a bounded function of δ, $c_n\overset{\triangle}{=}a_n+b_n\frac{\sigma^2_{\mathrm{B},X,\,i}(w_n)}{\sigma^2_{\mathrm{B},X,\,r}(w_n)}$ and $d_n^2\overset{\triangle}{=}\frac{2|\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w_n)|^2}{\sigma^2_{\mathrm{B},X,\,r}(w_n)}$. Using (51) and (29), we get

$$\begin{aligned}
&|p(y|w_n)\log(p_Y(y;\,F_X))|\\
&\qquad\leq M(\delta)\exp\left(-\frac{(y-c_n)^2}{d_n^2}\right)[\alpha+\beta' y^2].
\end{aligned}\tag{52}$$

Now, let us define $h(y)\overset{\triangle}{=}M(\delta)\exp\left(-\frac{(y-c_n)^2}{d_n^2}\right)[\alpha+\beta' y^2]$ for $y\in\mathbb{R}$. It is a straightforward task to verify that $\int_{\mathbb{R}} h(y)\,dy<\infty$. Hence, by the Dominated Convergence Theorem, $i_{\mathrm{B}}(w;\,F_X)$ is continuous over $\mathcal{D}$.

To show that the function $i_{\mathrm{B}}(w;\,F_X)$ is analytic over $\mathcal{D}$, it is sufficient to show that if $\oint_C i_{\mathrm{B}}(w;\,F_X)\,dw=0$, for any closed contour C in $\mathcal{D}$, then Morera's Theorem applies and it results in the analyticity of $i_{\mathrm{B}}(w;\,F_X)$ over $\mathcal{D}$. This contour integral

is given by

$$\oint_C i_{\mathrm{B}}(w; F_X)\,dw = \oint_C \int_{\mathbb{R}} p(y|w) \log\left(p_Y(y; F_X)\right)\,dy\,dw - \oint_C \frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w))\,dw. \quad (53)$$

Since $h(y)$ is finite, we define Γ_w as

$$\Gamma_w = \max_{w\in\mathcal{D}} \int_{\mathbb{R}} |p(y|w_n) \log\left(p_Y(y; F_X)\right)|\,dy, \quad (54)$$

and we can write

$$\left|\oint_C i_{\mathrm{B}}(w; F_X)\,dw\right| = \left|\oint_C \int_{\mathbb{R}} p(y|w) \log\left(p_Y(y; F_X)\right)\,dy\,dw\right| \le \oint_C \int_{\mathbb{R}} |p(y|w) \log\left(p_Y(y; F_X)\right)|\,dy\,dw \le \Gamma_w \ell_C < \infty, \quad (55)$$

where ℓ_C is the length of C which is finite as C is a closed curve. Therefore, by applying Fubini Theorem [13], one can change the order of integration in (53) and get

$$\oint_C i_{\mathrm{B}}(w; F_X)\,dw = \int_{\mathbb{R}} \log\left(p_Y(y; F_X)\right)\,dy \oint_C p(y|w)\,dw - \oint_C \frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w))\,dw. \quad (56)$$

It is clear that the complex functions $p(y|w)$ and $\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w)$ are analytic over $\mathcal{D}$. This implies that

$$\oint_C p(y|w)\,dw = 0, \quad (57)$$

$$\oint_C \frac{1}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 w))\,dw = 0, \quad (58)$$

which results in $\oint_C i_{\mathrm{B}}(w; F_X)\,dw = 0$ and thus by Morera's Theorem, $i_{\mathrm{B}}(w; F_X)$ is analytic over $\mathcal{D}$. Similarly, it can be shown that $i_{\mathrm{E}}(w; F_X)$ is also analytic over $\mathcal{D}$ and therefore, the equivocation density $r_{\mathrm{eq}}(w; F_X) = i_{\mathrm{B}}(w; F_X) - i_{\mathrm{E}}(w; F_X)$ is analytic over $\mathcal{D}$.

6) The Secrecy-Capacity-Achieving Input Distribution Is Discrete With a Finite Number of Mass Points: To prove the discreteness of the optimal input distribution F_X^*, we use a contradiction approach. To this end, let us assume that $\mathcal{S}_{F_X^*}$ has an infinite number of elements. In view of the optimality condition (45), the analyticity of $r_{\mathrm{eq}}(w; F_X)$ over $\mathcal{D}$ and the Identity Theorem of complex analysis along with the Bolzano-Weierstrass Theorem, if $S_{F_X^*}$ has an infinite number of mass points, we get $r_{\mathrm{eq}}(w; F_X^*) = C_{\mathrm{S}}$ for all $w \in \mathcal{D}$. Since $(-1/\eta_{\mathrm{B}}^2, +\infty) \subset \mathcal{D}$, any real variable $x \in (-1/\eta_{\mathrm{B}}^2, +\infty)$ also belongs to $\mathcal{D}$. This, in turn, implies that if $r_{\mathrm{eq}}(w; F_X) = C_{\mathrm{S}}$, $\forall\, w \in \mathcal{D}$, then

$$r_{\mathrm{eq}}(x; F_X^*) = C_{\mathrm{S}}, \quad \forall\, x \in (-1/\eta_{\mathrm{B}}^2, +\infty). \quad (59)$$

Next, we show that (59) results in a contradiction. By observing the bounds given in (25)–(30), one can easily show that

$$\int_{\mathbb{R}} \exp(-\alpha-\beta' y^2)[-\alpha-\beta' y^2]\,dy \le \int_{\mathbb{R}} p(y|x)\log(p_Y(y; F_X^*))\,dy \le \int_{\mathbb{R}} \exp(\alpha-\beta y^2)[\alpha+\beta' y^2]\,dy, \quad (60)$$

for all $x \in (-1/\eta_{\mathrm{B}}^2,\ A) \subset (-1/\eta_{\mathrm{B}}^2,\ +\infty)$. Similarly,

$$\int_{\mathbb{R}} \exp(-\mu-\xi' z^2)[-\mu-\xi' z^2]\,dz \le \int_{\mathbb{R}} p(z|x)\log(p_Z(z; F_X^*))\,dz \le \int_{\mathbb{R}} \exp(\mu-\xi z^2)[\mu+\xi' y^2]\,dz, \quad (61)$$

for all $x \in (-1/\eta_{\mathrm{B}}^2,\ A)$. Therefore, we can write

$$L \le -\int_{\mathbb{R}} p(y|x)\log(p_Y(y; F_X^*))\,dy + \int_{\mathbb{R}} p(z|x)\log(p_Z(y; F_X^*))\,dz \le U, \quad (62)$$

where the lower bound L and the upper bound U are given respectively as

$$L = \int_{\mathbb{R}} [-\mu-\xi' z^2]\exp(-\mu-\xi' z^2)\,dz + \int_{\mathbb{R}} [-\alpha-\beta' y^2]\,\exp(\alpha-\beta y^2)\,dy, \quad (63)$$

$$U = \int_{\mathbb{R}} [\mu+\xi' z^2]\exp(\mu-\xi z^2)\,dz + \int_{\mathbb{R}} [\alpha+\beta' y^2]\exp(-\alpha-\beta' y^2)\,dy. \quad (64)$$

Next, we establish that for finite positive reals β, β', ξ, ξ', μ and α, L and U are finite values. To prove the finiteness of L and U, it is sufficient to prove that L is finite as the finiteness of U follows along a similar line as that of L. We start by expanding L as $L = L_1 + L_2$, where L_1 and L_2 are given respectively as

$$L_1 = \int_{\mathbb{R}} [-\mu-\xi' z^2]\,\exp(-\mu-\xi' z^2)\,dz, \quad (65)$$

$$L_2 = \int_{\mathbb{R}} [-\alpha-\beta' y^2]\,\exp(\alpha-\beta y^2)\,dy. \quad (66)$$

Since the proof of the finiteness of L_2 is quite similar to that of L_1, we only show the finiteness of L_1. By direct integration, one can show that $L_1 = -\exp(-\mu)\,(\mu+\frac{1}{2})\sqrt{\frac{\pi}{\xi'}}$, which is a finite value for finite positive reals μ and ξ'. This implies that L is a finite value for finite positive reals μ, α, ξ', β and β'. Next, by substituting (21) and (22) into (59) and using the bounds in (63)–(64), we can write

$$L \le C_{\mathrm{S}} + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}\right) \le U. \quad (67)$$

Now, we define the sequence $\{x_n\}_{n\in\mathbb{N}}$ of distinct points in the interval $\mathbb{S} \triangleq (-1/\eta_{\mathrm{B}}^2, A)$ such that it is convergent to a limit point $x_0 = -1/\eta_{\mathrm{B}}^2$. We note that the limit point does not

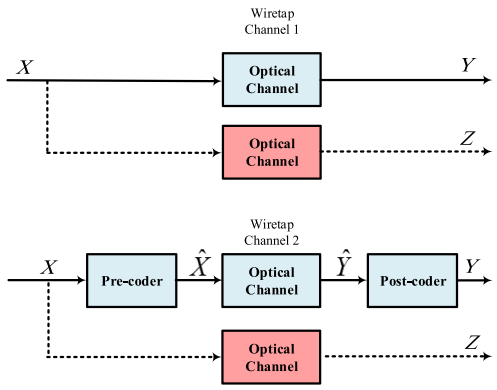

Fig. 3. Two optical wiretap channels with input-dependent Gaussian noise.

necessarily belong to $(-1/\eta_{\mathrm{B}}^2, A)$. Based on this, we have the following observations:

- x_n and $\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)$ are real values for all positive integers n.
- The limit $\lim_{n\to\infty}\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)$ exists and is equal to 0. This is established as follows:

$$\lim_{n\to\infty}\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)=\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2\lim_{n\to\infty}x_n)\overset{(a)}{=}0, \quad (68)$$

where (a) follows from the fact that $\lim_{n\to\infty}x_n = x_0 = -1/\eta_{\mathrm{B}}^2$ and $(1+\eta_{\mathrm{B}}^2\ x_0)=0$.

Following the results in [4, Theorem 3] and using (67), we can write

$$\begin{aligned}\lim_{n\to\infty}(L-C_{\mathrm{S}}) &\le \lim_{n\to\infty}\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right)\\ &\le \lim_{n\to\infty}(U-C_{\mathrm{S}}). \end{aligned}\quad (69)$$

We note that $\lim_{n\to\infty}\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right)=-\infty$ (as $\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2\ x_0)$ is a positive finite value due to the fact that $\eta_{\mathrm{B}}^2>\eta_{\mathrm{E}}^2$) and the $\lim_{n\to\infty}(L-C_{\mathrm{S}})=L-C_{\mathrm{S}}$ is a finite value,[1] thus a contradiction occurs. This, in turn, implies that the support set $\mathcal{S}_{F_X^*}$ cannot have an infinite number of elements and therefore, the optimal input distribution F_X^* is discrete with a finite number of mass points.

C. Proof of Proposition 1

Suppose, to the contrary, that $x=0$ does not belong to the support set of the optimal input distribution $\mathcal{S}_{F_X^*}$. Let $0<x_1\le x_2\le\ldots\le x_N$ be the mass points in the set $\mathcal{S}_{F_X^*}$. Consider two optical wiretap channels with input-dependent Gaussian noises depicted in Fig. 3. Wiretap channel 1 is the original optical wiretap channel, and wiretap channel 2 is obtained from wiretap channel 1 by appending a pre-coder and a post-coder before and after the inner optical channel in the legitimate user's link. Specifically, $\hat{X}=X-x_1$ and $Y=\hat{Y}+\hat{N}_{\mathrm{B}}$, where $\hat{N}_{\mathrm{B}}$ is an additive Gaussian noise with mean x_1 and variance $x_1\sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2$ and is independent from $\hat{Y}$. For any $x\ge x_1$, the conditional probability density functions $p(y|x)$ and $p(z|x)$ are the same in both wiretap channels. Thus, the joint probability density functions of $p(y,x)$ and $p(z,x)$ in the two wiretap channels are also the same, if the input distribution is F_X^*. As a result, C_{S} is identical in both wiretap channels.

In the second wiretap channel, as $X,\hat{X},\hat{Y},Y$ and Z form the Markov chain $X\to\hat{X}\to\hat{Y}\to Y\to Z$, we have $I(\hat{X};\hat{Y}|Z)\ge I(X;Y|Z)$ by the data processing inequality. This indicates that $I(\hat{X};\hat{Y})-I(\hat{X};Z)\ge I(X;Y)-I(X;Z)$. Now, let $F_{\hat{X}}^*$ be the distribution function of $\hat{X}$ when the distribution function of X is F_X^*. Clearly, $F_{\hat{X}}^*$ satisfies the nonnegativity and peak-intensity constraints. Hence, $F_{\hat{X}}^*$ is also secrecy-capacity-achieving for wiretap channel 1. Based on Lemma 1, the secrecy-capacity-achieving output distribution is unique, as a result, $p_Y(y;F_X^*)=p_Y(y;F_{\hat{X}}^*)$. Therefore, for wiretap channel 2, given the input distribution function of X is F_X^*, the probability density functions for Y and $\hat{Y}$ are the same, which is not possible since $\mathbb{E}[Y]=\mathbb{E}[\hat{Y}]+x_1$. Hence, we reach a contradiction and the proposition follows.

D. Proof of Theorem 2

This section presents the proof of Theorem 2 by extending the analysis in the previous section to the entire rate-equivocation region. This extension entails generalizing the contradiction argument in the proof of Theorem 1 to the case when an additional mutual information term is present in the objective function. We start by noting that the objective function $g_\lambda(F_X)$ in (15) is strictly concave, and the feasible set $\mathcal{A}^+$ is compact and convex, therefore, the optimization problem in (15) has a unique maximizer. We denote the optimal input distribution for (15) as F_X^* which depends on the value λ.

Now, we obtain the KKT optimality conditions for the optimal input distribution of the optimization problem in (15). Since the objective function $g_\lambda(F_X)$ is weakly differentiable, we have

$$\lim_{\theta\to0}\frac{g_\lambda((1-\theta)\,F_{X_0}+\theta F_X)-g_\lambda(F_{X_0})}{\theta} \quad (70)$$

$$=\int_0^A[\lambda\, i_{\mathrm{B}}(x;F_{X_0})+(1-\lambda)\,r_{\mathrm{eq}}(x;F_X)]\,dF_X(x)-g_\lambda(F_{X_0}). \quad (71)$$

Following similar steps mentioned in the proof of Theorem 1, the KKT optimality conditions for the optimality of F_X^* are obtained as follows

$$\begin{aligned}&\lambda\, i_{\mathrm{B}}(x;F_X^*)+(1-\lambda)\,r_{\mathrm{eq}}(x;F_X^*)\le\lambda I_{\mathrm{B}}(F_X^*)\\ &\qquad+(1-\lambda)(I_{\mathrm{B}}(F_X^*)-I_{\mathrm{E}}(F_X^*)),\quad \forall x\in[0,A], \end{aligned}\quad (72)$$

$$\begin{aligned}&\lambda\, i_{\mathrm{B}}(x;F_X^*)+(1-\lambda)\,r_{\mathrm{eq}}(x;F_X^*)=\lambda\, I_{\mathrm{B}}(F_X^*)\\ &\qquad+(1-\lambda)(I_{\mathrm{B}}(F_X^*)-I_{\mathrm{E}}(F_X^*)),\quad \forall x\in\mathcal{S}_{F_X^*}. \end{aligned}\quad (73)$$

Next, we show that the optimal input distribution F_X^* has a finite support. To this end, we use similar steps mentioned in the proof of Theorem 1 and prove the discreteness of F_X^* by a contradiction approach and using the optimality conditions in (72)–(73).

[1] It is shown in Section V.B that $C_{\mathrm{S}}\le\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{E}}^2}{\sigma_{\mathrm{E}}^2}\right)$. Furthermore, as $C_{\mathrm{S}}\ge 0$, we have $0\le C_{\mathrm{S}}\le\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{E}}^2}{\sigma_{\mathrm{E}}^2}\right)$. This implies that C_{S} is a finite value.

Let us assume that $\mathcal{S}_{F_X^*}$ has an infinite number of elements. Under such an assumption, (73), the analyticity of $i_{\mathrm{B}}(w; F_X^*)$ and $i_{\mathrm{E}}(w; F_X^*)$ over $\mathcal{D}$ in the complex plane and the Identity Theorem of complex analysis imply that

$$\lambda\, i_{\mathrm{B}}(x; F_X^*) + (1-\lambda)\, r_{\mathrm{eq}}(x; F_X^*) = \lambda\, I_{\mathrm{B}}(F_X^*) + (1-\lambda)(I_{\mathrm{B}}(F_X^*) - I_{\mathrm{E}}(F_X^*)), \quad \forall\, x \in (-1/\eta_{\mathrm{B}}^2, +\infty). \tag{74}$$

Next, we show that (74) results in a contradiction. To this end, by using (60) and (61) and the fact that $(1-\lambda)$ is nonnegative for all $\lambda \in [0, 1]$, we can bound (74) as

$$\widetilde{L} \le I_{\mathrm{B}}(F_X^*) - (1-\lambda)\, I_{\mathrm{E}}(F_X^*) + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}\right) + \frac{\lambda}{2}\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)) \le \widetilde{U}, \tag{75}$$

where $\widetilde{L}$ and $\widetilde{U}$ are given by

$$\widetilde{L} = (1-\lambda)\int_{\mathbb{R}} [-\mu - \xi' z^2]\, \exp(-\mu - \xi' z^2)\, dz + \int_{\mathbb{R}} [-\alpha - \beta' y^2]\, \exp(\alpha - \beta y^2)\, dy, \tag{76}$$

$$\widetilde{U} = (1-\lambda)\int_{\mathbb{R}} [\mu + \xi' z^2]\, \exp(\mu - \xi z^2)\, dz + \int_{\mathbb{R}} [\alpha + \beta' y^2]\, \exp(-\alpha - \beta' y^2)\, dy. \tag{77}$$

Invoking similar arguments for the proving the finiteness of L and U given in (63)–(64), one can show that the lower bound $\widetilde{L}$ and the upper bound $\widetilde{U}$ are also finite values. Now, let $\{x_n\}_{n\in\mathbb{N}}$ be a convergent sequence of distinct points in $\mathbb{S}$ with a limit point $x_0 = -1/\eta_{\mathrm{B}}^2$. It is clear that 1) x_n and $\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)$ are real for all positive integers n; 2) $\lim_{n\to\infty} \sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n) = 0$. Following the results in [4, Theorem 3] and using (75), we get

$$\begin{aligned}
&\lim_{n\to\infty}\left[\widetilde{L} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)\right] \\
&\le \lim_{n\to\infty} \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right) + \frac{\lambda}{2}\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)) \\
&\le \lim_{n\to\infty}\left[\widetilde{U} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)\right],
\end{aligned} \tag{78}$$

We note that $\lim_{n\to\infty} \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right) = -\infty$, while $\frac{\lambda}{2}\log(2\pi e\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{E}}^2 x_0))$ and $\widetilde{L} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)$ are finite values. Hence, we reach a contradiction; implying that the optimal input distribution F_X^* has a finite support.

V. Asymptotic Results for a Peak-Intensity Constraint

This section provides the asymptotic analysis on the secrecy capacity of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints. First, the secrecy capacity is investigated for asymptotically small values of A. Second, we prove that for high-intensity regime, the secrecy capacity can be bounded by a constant value implying that it does not scale with the peak-intensity constraint in this regime.

A. Low-Intensity Results

For relatively small values of the peak-intensity constraint A, we use the results shown in [14] and we can write

$$I(X;Y) - I(X;Z) = \frac{1}{2}[J_{\mathrm{B}}(0) - J_{\mathrm{E}}(0)]\,\mathrm{Var}(X) + o(A^2), \tag{79}$$

where $o(A^2)$ denotes a term that tends to 0 faster than A^2, $\mathrm{Var}(X)$ is the variance of the random variable X, $J_{\mathrm{B}}(0)$ and $J_{\mathrm{E}}(0)$ denote the Fisher information of the legitimate user's and wiretap channel at 0 and $J(x)$ is given by

$$J(x) = \int_y \frac{(\frac{d}{dx}p(y|x))^2}{p(y|x)}\, dy. \tag{80}$$

For the channel laws (18) and (19), we have

$$J_{\mathrm{B}}(x) = \frac{2+\eta_{\mathrm{B}}^4\sigma_{\mathrm{B}}^2 + 2\eta_{\mathrm{B}}^2 x}{2\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)^2}, \tag{81}$$

$$J_{\mathrm{E}}(x) = \frac{2+\eta_{\mathrm{E}}^4\sigma_{\mathrm{E}}^2 + 2\eta_{\mathrm{E}}^2 x}{2\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)^2}, \tag{82}$$

such that

$$J_{\mathrm{B}}(0) = \frac{2+\eta_{\mathrm{B}}^4\sigma_{\mathrm{B}}^2}{2\sigma_{\mathrm{B}}^2}, \tag{83}$$

$$J_{\mathrm{E}}(0) = \frac{2+\eta_{\mathrm{E}}^4\sigma_{\mathrm{E}}^2}{2\sigma_{\mathrm{E}}^2}. \tag{84}$$

Therefore, for small values of A, the secrecy capacity is

$$C_{\mathrm{S}} = \frac{1}{2}[J_{\mathrm{B}}(0) - J_{\mathrm{E}}(0)] \max_{F_X\in\mathcal{A}^+} \mathrm{Var}(X) + o(A^2). \tag{85}$$

Proposition 2: In the regime $A \ll 1$, the secrecy capacity under the peak-intensity constraint is as follows

$$C_{\mathrm{S}}(A) = \frac{A^2}{8}\left(\frac{1}{\sigma_{\mathrm{B}}^2} - \frac{1}{\sigma_{\mathrm{E}}^2} + \frac{1}{2}(\eta_{\mathrm{B}}^4 - \eta_{\mathrm{E}}^4)\right) + o(A^2). \tag{86}$$

Proof: See Appendix B. ∎

Proposition 2 indicates the secrecy capacity to be a quadratic function of the peak-intensity constraint A in the low-intensity regime. Furthermore, as we show in Appendix B, under constraint (5) two mass points located at 0 and A with equal probabilities are optimal in this regime.

B. High-Intensity Results

In this section, we provide an upper bound on the secrecy capacity that holds for any value of A. Based on (46), the secrecy capacity can be simplified as

$$C_{\mathrm{S}} = h_Y(F_X^*) - h_Z(F_X^*) + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{E}}^2}{\sigma_{\mathrm{B}}^2}\right) + \frac{1}{2}\mathbb{E}_{F_X^*}\left[\log\left(\frac{1+\eta_{\mathrm{E}}^2 x}{1+\eta_{\mathrm{B}}^2 x}\right)\right]. \tag{87}$$

Since the optical wiretap channel with input-dependent Gaussian noise is stochastically degraded and based on the fact that $\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 = \sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2$ and $\sigma_{\mathrm{E}}^2 > \sigma_{\mathrm{B}}^2$, one can write $Z = Y + N_{\mathrm{D}}$ for some zero-mean Gaussian random variable N_D

with variance $\sigma_{\mathrm{D}}^2 = \sigma_{\mathrm{E}}^2 - \sigma_{\mathrm{B}}^2$. Therefore, $h(Z) > h(Z|N_{\mathrm{D}}) = h(Y)$ and consequently $h(Z) > h(Y)$ for any nontrivial input distribution F_X^*. Furthermore, as $\eta_{\mathrm{E}}^2 < \eta_{\mathrm{B}}^2$ and $x \geq 0$, we have $\frac{1}{2}\mathbb{E}_{F_X^*}\left[\log\left(\frac{1+\eta_{\mathrm{E}}^2\, x}{1+\eta_{\mathrm{B}}^2\, x}\right)\right] \leq 0$. As a result, $C_{\mathrm{S}} \leq \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{E}}^2}{\sigma_{\mathrm{B}}^2}\right)$ for any value of A. This, in turn, implies that the secrecy capacity in the regime $A \to \infty$ does not scale with the peak-intensity constraint A and converges to a real and positive constant, i.e.,

$$C_{\mathrm{S}}(A) = O(1), \tag{88}$$

where $O(1)$ is a function such that for large enough A, the secrecy capacity is at most k_0, for some real number $k_0 > 0$.

VI. The Case of Peak- and Average-Intensity Constraints

In this section, we generalize the discreteness result for the optimal input distribution when an additional average intensity constraint is imposed on the channel input. We begin with generalizing Theorems 1 and 2 for the case when we have both peak- and average-intensity constraints by establishing parallels to the proof of these theorems. Let the average intensity constraint be P. The new feasible set for the input distribution is

$$\mathcal{M}^+ \triangleq \left\{F_X : \int_0^A dF_X(x) = 1, \quad \int_0^A x\, dF_X(x) \leq P\right\}. \tag{89}$$

We first consider the secrecy capacity

$$C_{\mathrm{S}} = \max_{F_X \in \mathcal{M}^+} \{I(X;Y) - I(X;Z)\}. \tag{90}$$

Similar to the lines provided for the proof of Theorem 1, here too, the mutual information difference $I(X;Y) - I(X;Z)$ is strictly concave and continuous in the input distribution F_X. Furthermore, $\mathcal{M}^+$ is a compact and convex set [12, Appendix A.1]. Thus the necessary and sufficient conditions in (44)–(45) take the new form

$$r_{\mathrm{eq}}(x; F_X^*) - \gamma\, x \leq C_{\mathrm{S}} - \gamma\, \mathbb{E}[X], \quad \forall x \in [0, A], \tag{91}$$

$$r_{\mathrm{eq}}(x; F_X^*) - \gamma\, x = C_{\mathrm{S}} - \gamma\, \mathbb{E}[X], \quad \forall x \in \mathcal{S}_{F_X^*}, \tag{92}$$

$$\gamma\ (\mathbb{E}[X] - P) = 0, \tag{93}$$

for some $\gamma \geq 0$. Assuming that the average intensity constraint is tight, we have $\gamma > 0$. For the case of $\gamma = 0$, the only imposed constraints are the nonnegativity and peak-intensity constraints and we have already proven in Theorem 1 that the optimal input distribution is discrete. Hence, (91)–(93) can be rewritten as

$$r_{\mathrm{eq}}(x; F_X^*) - \gamma\, x \leq C_{\mathrm{S}} - \gamma\, \mathbb{E}[X], \quad \forall x \in [0, A], \tag{94}$$

$$r_{\mathrm{eq}}(x; F_X^*) - \gamma\, x = C_{\mathrm{S}} - \gamma\, \mathbb{E}[X], \quad \forall x \in \mathcal{S}_{F_X^*}, \tag{95}$$

$$\mathbb{E}[X] = P. \tag{96}$$

Next, we prove by contradiction that the optimal input distribution F_X^* satisfying (94)–(96) must be discrete with a finite number of mass points. Now, let us assume that $\mathcal{S}_{F_X^*}$ has an infinite number of elements. In view of the optimality condition (94)–(96), the analyticity of $r_{\mathrm{eq}}(w; F_X)$ and w over $\mathcal{D}$ and the Identity Theorem of complex analysis, along with the Bolzano-Weierstrass Theorem, if $\mathcal{S}_{F_X^*}$ has an infinite number of mass points, we get $r_{\mathrm{eq}}(w; F_X^*) - \gamma\ w = C_{\mathrm{S}} - \gamma\ P$ for all $w \in \mathcal{D}$, which results in

$$r_{\mathrm{eq}}(x; F_X^*) - \gamma\ x = C_{\mathrm{S}} - \gamma\ P, \quad x \in (-1/\eta_{\mathrm{B}}^2, +\infty), \tag{97}$$

$$\mathbb{E}[X] = P. \tag{98}$$

Using the bounds given in (63)–(64), one can write

$$L \leq C_{\mathrm{S}} - \gamma\ P + \gamma\ x + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}\right) \leq U. \tag{99}$$

Defining a convergent sequence of distinct points $\{x_n\}_{n\in\mathbb{N}}$ in $\mathbb{S}$ with a limit point $x_0 = -1/\eta_{\mathrm{B}}^2$, we have: 1) x_n and $\sigma_{\mathrm{B}}^2(1 + \eta_{\mathrm{B}}^2 x_n)$ are real for all positive integers n; 2) $\lim_{n\to\infty} \sigma_{\mathrm{B}}^2(1 + \eta_{\mathrm{B}}^2 x_n) = 0$. Following the results in [4, Theorem 3] and using (99) we can write

$$\begin{aligned}
&\lim_{n\to\infty} (L - C_{\mathrm{S}}) \\
&\quad \leq \lim_{n\to\infty} \left[\gamma\ x_n - \gamma\ P + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right)\right] \\
&\quad \leq \lim_{n\to\infty} (U - C_{\mathrm{S}}).
\end{aligned} \tag{100}$$

Since $\lim_{n\to\infty} \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right) = -\infty$ and $L - C_{\mathrm{S}}$ is a finite value, a contradiction occurs. This, in turn, implies that the support set $\mathcal{S}_{F_X^*}$ cannot have an infinite number of mass points and therefore the optimal input distribution F_X^* under the nonnegativity, peak- and average-intensity constraints, is also discrete with a finite number of mass points. Additionally, following along similar lines as in Proposition 1, one can prove that $x = 0$ belongs to the support set $\mathcal{S}_{F_X^*}$ of the secrecy-capacity-achieving input distribution with peak- and average-intensity constraints.

Finally, we extend this contradiction argument to the entire rate-equivocation region. Consider the optimization problem for determining the boundary point of the rate-equivocation region

$$\max_{F_X \in \mathcal{M}^+} \{\lambda\, I(X;Y) + (1-\lambda)\,[I(X;Y) - I(X;Z)]\}. \tag{101}$$

We note that if the average intensity constraint is not tight, i.e., $\mathbb{E}[X] < P$, the problem reduces to the case where only the nonnegativity and peak-intensity constraints are present, in which case the optimal input distribution is discrete with a finite support by Theorem 2. Hence, without loss of generality, we assume that the average intensity constraint is tight and the necessary and sufficient optimality conditions for (101) are

$$\begin{aligned}
&\lambda\, i_{\mathrm{B}}(x; F_X^*) + (1-\lambda)\, r_{\mathrm{eq}}(x; F_X^*) - \gamma\ x \\
&\qquad \leq \lambda\, I_{\mathrm{B}}(F_X^*) + (1-\lambda)[I_{\mathrm{B}}(F_X^*) - I_{\mathrm{E}}(F_X^*)] - \gamma\ P, \\
&\qquad\qquad \forall x \in [0, A],
\end{aligned} \tag{102}$$

$$\begin{aligned}
&\lambda\, i_{\mathrm{B}}(x; F_X^*) + (1-\lambda)\, r_{\mathrm{eq}}(x; F_X^*) - \gamma\ x \\
&\qquad = \lambda\, I_{\mathrm{B}}(F_X^*) + (1-\lambda)[I_{\mathrm{B}}(F_X^*) - I_{\mathrm{E}}(F_X^*)] - \gamma\ P, \\
&\qquad\qquad \forall x \in \mathcal{S}_{F_X^*},
\end{aligned} \tag{103}$$

$$\mathbb{E}[X] = P. \tag{104}$$

Next, we prove by contradiction that the input distribution F_X^* satisfying (102)–(104) must be a discrete distribution with

a finite support. Assume, on the contrary, that $\mathcal{S}_{F_X^*}$ has an infinite number of elements. In view of (102)–(104) and the analyticity of $i_{\mathrm{B}}(w; F_X)$, $r_{\mathrm{eq}}(w; F_X)$ and w over $\mathcal{D}$ and the Identity Theorem of complex analysis, we have

$$\begin{aligned}\lambda\, i_{\mathrm{B}}(x; F_X^*) + (1-\lambda)\, r_{\mathrm{eq}}(x; F_X^*) - \gamma\, x \\ = \lambda\, I_{\mathrm{B}}(F_X^*) + (1-\lambda)[I_{\mathrm{B}}(F_X^*) - I_{\mathrm{E}}(F_X^*)] - \gamma\, P, \\ \forall x \in (-1/\eta_{\mathrm{B}}^2, +\infty), \quad (105)\end{aligned}$$

$$\mathbb{E}[X] = P. \quad (106)$$

Using the bounds presented in (76)–(77), one can verify that

$$\begin{aligned}\widetilde{L} \le I_{\mathrm{B}}(F_X^*) - (1-\lambda)\, I_{\mathrm{E}}(F_X^*) + \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}\right) \\ + \frac{\lambda}{2}\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)) + \gamma\,(x-P) \le \widetilde{U}. \quad (107)\end{aligned}$$

Now, let $\{x_n\}_{n\in\mathbb{N}}$ be a convergent sequence of distinct points in $\mathbb{S}$ such that it is converging to a limit point $x_0 = -1/\eta_{\mathrm{B}}^2$. Based on this, we have the following results: 1) x_n and $\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)$ are real for all positive integers n; 2) $\lim_{n\to\infty}\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n) = 0$. Following the results in [4, Theorem 3] and using (107), we get

$$\begin{aligned}&\lim_{n\to\infty}[\widetilde{L} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)] \\ &\le \lim_{n\to\infty}\left[\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right) + \frac{\lambda}{2}\log(2\pi e\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)) \right.\\ &\quad \left. + \gamma\,(x_n - P)\right] \\ &\le \lim_{n\to\infty}[\widetilde{U} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)]. \quad (108)\end{aligned}$$

Since, $\lim_{n\to\infty}\frac{1}{2}\log\left(\frac{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x_n)}{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x_n)}\right) = -\infty$ while $\widetilde{L} - I_{\mathrm{B}}(F_X^*) + (1-\lambda)\, I_{\mathrm{E}}(F_X^*)$ is a finite value, a contradiction occurs and we conclude that the support set $\mathcal{S}_{F_X^*}$ has a finite number of mass points.

VII. Asymptotic Results for the Secrecy Capacity Under Peak- and Average-Intensity Constraints

This section provides the asymptotic analysis on the secrecy capacity of the optical wiretap channel with input-dependent Gaussian noise under peak- and average-intensity constraints. First, the secrecy capacity is investigated for asymptotically small values of A and P with their ratio $\kappa \triangleq \frac{P}{A} \in (0, \frac{1}{2})$. Second, an upper bound on the secrecy capacity will be given that holds true for any value of A and κ.

A. Low-Intensity Results

For the small values of A and κ, based on the results in [14] we can write

$$C_{\mathrm{S}} = \frac{1}{2}[J_{\mathrm{B}}(0) - J_{\mathrm{E}}(0)] \max_{F_X\in\mathcal{M}^+} \mathrm{Var}(X) + o(A^2). \quad (109)$$

Proposition 3: In the regime $A \ll 1$ and $\kappa \in (0, \frac{1}{2})$, the secrecy capacity under the peak- and average-intensity constraints is as follows

$$C_{\mathrm{S}}(A,\kappa) = \frac{A^2}{2}\kappa(1-\kappa)\left(\frac{1}{\sigma_{\mathrm{B}}^2} - \frac{1}{\sigma_{\mathrm{E}}^2} + \frac{1}{2}(\eta_{\mathrm{B}}^4 - \eta_{\mathrm{E}}^4)\right) + o(A^2). \quad (110)$$

Proof: See Appendix C. ∎

Similar to the case with a peak-intensity constraint, Proposition 3 reflects that the secrecy capacity under peak- and average-intensity constraints scales quadratically in A. Furthermore, as we show in Appendix C, the secrecy-capacity-achieving input distribution possesses two mass points located at 0 and A with probabilities κ and $1-\kappa$, respectively. Additionally, we note that based on Appendix C, the average intensity constraint is inactive when $\kappa \in [\frac{1}{2}, 1]$ and consequently $C_{\mathrm{S}}(A,\kappa)$ is given by (86).

B. High-Intensity Results

In the regime $A \to \infty$ and $P \to \infty$ with their ratio κ kept fixed, the secrecy capacity $C_{\mathrm{S}} \le \frac{1}{2}\log\left(\frac{\sigma_{\mathrm{E}}^2}{\sigma_{\mathrm{B}}^2}\right)$ and the proof follows along similar lines as in Proposition 2. This implies that the secrecy capacity under peak- and average-intensity constraints in the high-intensity regime does not scale with the constraints and therefore converges to a real and positive constant.

VIII. Numerical Results

In this section, we provide numerical results for the secrecy capacity and the entire rate-equivocation region of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity, peak- and average-intensity constraints.

Figure 4 provides a plot of the equivocation density for an optimal input distribution for $A = 4$, $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, and $\eta_{\mathrm{E}}^2 = 0.125$. We numerically found that for these parameters, the optimal input distribution is ternary with mass points located at $x = 0$, 2.025 and 4 with probability masses 0.2862, 0.3045 and 0.4093, respectively. We observe that $C_{\mathrm{S}} - r_{\mathrm{eq}}(x; F_X)$ is generally nonnegative and is equal to zero at the optimal mass points; verifying the optimality conditions in (44) and (45).

Figure 5 illustrates the secrecy capacity C_{S} and the difference $C_{\mathrm{B}} - C_{\mathrm{E}}$ versus the peak-intensity constraint A, where C_{B} and C_{E} are the legitimate user's and the eavesdropper's

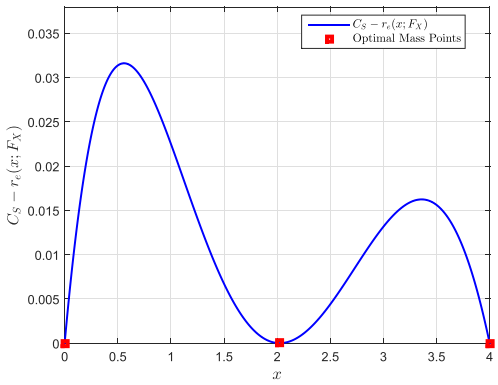

Fig. 4. Illustration of $C_{\mathrm{S}} - r_{\mathrm{eq}}(x; F_X)$ yielded by the optimal input distribution when $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, $\eta_{\mathrm{E}}^2 = 0.125$ and $A = 4$.

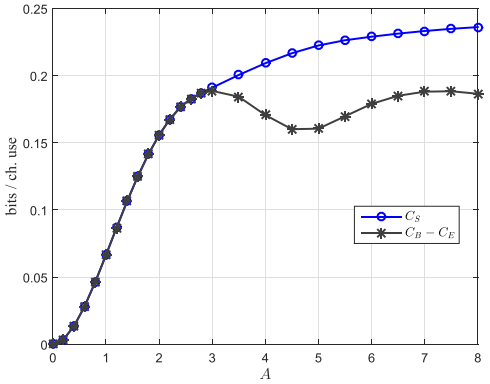

Fig. 5. The secrecy capacity for $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, and $\eta_{\mathrm{E}}^2 = 0.125$ versus the peak-intensity constraint A.

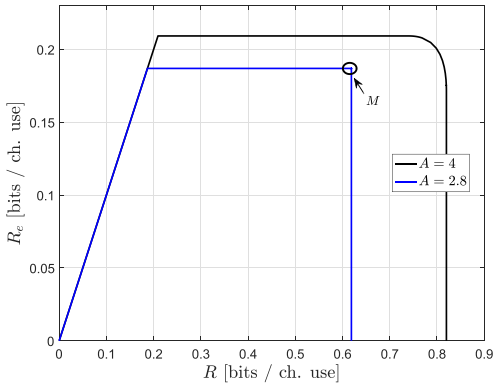

Fig. 6. The rate-equivocation region for $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, and $\eta_{\mathrm{E}}^2 = 0.125$ under peak-intensity constraints $A = 2.8$ and $A = 4$. Point M refers to the case when secrecy capacity and capacity are achieved simultaneously.

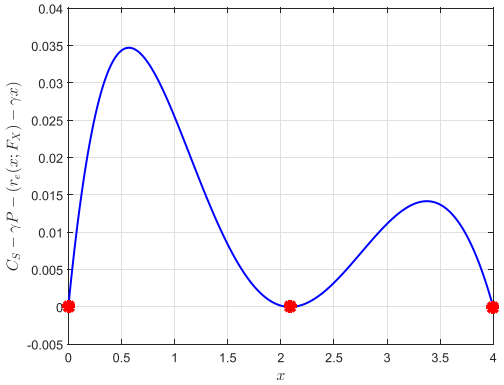

Fig. 7. Illustration of $C_{\mathrm{S}} - r_{\mathrm{eq}}(x; F_X) + \gamma\ (x - P)$ yielded by the optimal input distribution for $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, $\eta_{\mathrm{E}}^2 = 0.125$, $A = 4$ and $\kappa = 0.375$. The corresponding Lagrangian multiplier is 0.0187.

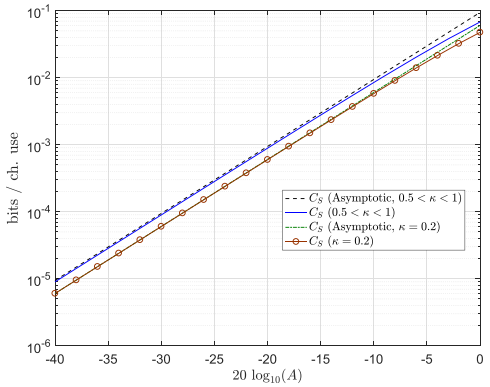

Fig. 8. The asymptotic and exact secrecy capacity for $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, and $\eta_{\mathrm{E}}^2 = 0.125$ versus A for both peak- and average-intensity constraints.

capacities, respectively. We observe that this difference is in general a lower bound for the secrecy capacity C_{S} which can be easily proven. We also observe that, for small values of A, $C_{\mathrm{B}} - C_{\mathrm{E}}$ and C_{S} are identical. However, as A increases, $C_{\mathrm{B}} - C_{\mathrm{E}}$ and C_{S} become different. Similar to the secrecy capacity results of the Gaussian wiretap channel under a peak-power constraint provided in [9], here too, $I(X;Y)$ and $I(X;Z)$ are maximized by the same discrete distribution, however, $I(X;Y) - I(X;Z)$ is maximized by a different distribution. As a specific example, when $A = 4$, while both $I(X;Y)$ and $I(X;Z)$ are maximized by the same *binary* distribution with mass points at $x = 0$ and 4 with probability masses 0.5088 and 0.4912, respectively, $I(X;Y) - I(X;Z)$ is maximized by a *ternary* distribution with mass points at $x = 0$, 2.025 and 4 with probability masses 0.2862, 0.3045 and 0.4093, respectively. This explains the difference between C_{S} and $C_{\mathrm{B}} - C_{\mathrm{E}}$ at $A = 4$ in this figure.

Figure 6 depicts the entire rate-equivocation region of the optical wiretap channel with input-dependent Gaussian noise under nonnegativity and peak-intensity constraints when $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, and $\eta_{\mathrm{E}}^2 = 0.125$ for two different values of A. When $A = 2.8$, it is clear from the figure that both the secrecy capacity and the capacity can be attained simultaneously (Point "M" in the figure). In particular, for $A = 2.8$, the binary input distribution with mass points located at $x = 0$ and 2.8 with probabilities 0.5183 and 0.4817, respectively, achieves both the capacity and the secrecy capacity. This implies that, when $A = 2.8$, the transmitter can communicate with the legitimate user at the capacity while achieving the maximum equivocation at the eavesdropper. On the other hand, when $A = 4$, the secrecy capacity and the capacity cannot be achieved simultaneously (notice the curved shape in the figure). More specifically, for $A = 4$, the binary input distribution with mass points located at $x = 0$ and 4 with probabilities 0.5088 and 0.4912 achieves the capacity, while a ternary distribution with mass points located at $x = 0$, 2.025, 4 with probability masses 0.2862, 0.3045 and 0.4093, respectively, 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 its equivocation in the sense that, to increase the communication rate, one must compromise from the equivocation of this communication, and to increase the achieved equivocation, one must compromise from the communication rate.

Figure 7 provides illustrations for the effect of the average intensity constraint on the secrecy-capacity-achieving input distribution for $\sigma_{\mathrm{B}}^2 = 1$, $\sigma_{\mathrm{E}}^2 = 2$, $\eta_{\mathrm{B}}^2 = 0.25$, $\eta_{\mathrm{E}}^2 = 0.125$ and $A = 4$. We observe that for $\kappa > \frac{1}{2}$, the average intensity constraint is inactive. In this case, in view of Theorem 1, the optimal input distribution is a ternary distribution with

mass points at $x = 0$, 2.025 and 4 with probability masses 0.2862, 0.3045 and 0.4093, respectively. We now impose an average intensity constraint with corresponding $\kappa = 0.375$. For this case, a ternary input distribution with mass points located at $x = 0$, 2.0869 and 4 with probability masses 0.4770, 0.3093 and 0.2136, respectively, is optimal and the corresponding Lagrangian multiplier is 0.0187.

Finally, in Fig. 8, we plot the exact and asymptotic secrecy capacity results versus the peak-intensity constraint A for both the peak- and average-intensity constraints in the low-intensity regime. From the figure, we observe that our asymptotic results for the secrecy capacity given in (86) and (110) are in precise agreement with the numerical results. Furthermore, we observe that imposing the average intensity constraint in addition to the peak-intensity constraint reduces the secrecy capacity.

IX. Conclusion

This paper studies the optical wiretap channel with input-dependent Gaussian noise under nonnegativity, peak- and average-intensity constraints. It is shown that the secrecy capacity and the boundary of the entire rate-equivocation region are achieved by discrete input distributions with finite supports. Furthermore, under such constraints, the optimal input distribution always possesses a mass point at the origin. An interesting result that this paper reveals is that similar to the case for the Gaussian wiretap channel with a peak-power constraint, here too, we observe that under nonnegativity and peak-intensity constraints, the secrecy capacity and the capacity cannot be obtained simultaneously in general, i.e., there is a tradeoff between the rate and its equivocation in the sense that, to increase the communication rate, one must compromise from the equivocation rate, and conversely to increase the achieved equivocation rate, one must compromise from the communication rate. We extend the discreteness result for the case when we have both peak- and average-intensity constraints. Finally, we provide an asymptotic analysis on the secrecy capacity. It is shown that in the low-intensity regime, the secrecy capacity scales quadratically with the peak-intensity constraint, while in the high-intensity regime, it is upper-bounded by a finite value implying that it does not scale with the constraints.

Appendix A
Proof of Lemma 1

We start the proof by noting that for random variables X, Y and Z that form the Markov chain $X \to Y \to Z$, $I(X;Y|Z)$ is a concave function in F_X [15, Appendix A]. Now, let X_1 and X_2 be two channel inputs generated by F_{X_1} and F_{X_2}, respectively, and Q be a binary-valued random variable such that

$$p(y,z,x|q) = \begin{cases} p(y,z|x)\, p_{X_1}(x), & q = 1, \\ p(y,z|x)\, p_{X_2}(x), & q = 2, \end{cases} \tag{111}$$

where $p_{X_1}(x)$ and $p_{X_2}(x)$ be the probability density functions (PDF) of the random variables X_1 and X_2. Based on (111), we have the following Markov chain

$$Q \to X \to Y \to Z. \tag{112}$$

Following along the same lines as [15, Appendix A], one can show that

$$I(X;Y|Z,Q) - I(X;Y|Z) = -I(Q;Y|Z). \tag{113}$$

Since $I(Q;Y|Z) \geq 0$, $I(X;Y|Z,Q) \leq I(X;Y|Z)$. This implies that $I(X;Y|Z)$ is a concave function in F_X. Now, we prove that with the Markov chain $Q \to X \to Y \to Z$, $I(X;Y|Z)$ is strictly concave in F_X, i.e., $I(Q;Y|Z) > 0$. Assume, to the contrary, that $I(Q;Y|Z) = 0$. This implies that random variables Q, Y and Z also form the Markov chain

$$Q \to Z \to Y. \tag{114}$$

Furthermore, from the Markov chain (112), we have

$$Q \to X \to Z. \tag{115}$$

Combining Markov chains (114) and (115) results in a new Markov chain given by

$$Q \to X \to Z \to Y. \tag{116}$$

Now, based on (112) and (116), we obtain the following

$$\begin{aligned} p(y,z,x)\big|_{\text{Markov chain (112)}} &= p(y,z,x)\big|_{\text{Markov chain (116)}} \\ p_X(x)\, p(y|x)\, p(z|y) &= p_X(x)\, p(z|x)\, p(y|z) \\ \frac{p(y|x)}{p(z|x)} &= \frac{p(y|z)}{p(z|y)}. \end{aligned} \tag{117}$$

We note that (117) holds for any $y, z \in \mathbb{R}$ and $x \in \mathcal{S}_{F_X}$, where $\mathcal{S}_{F_X}$ is the support set of F_X. As a result, for fixed values of y and z the right hand side (RHS) of (118) is fixed, while the left hand side (LHS) is a function of x. Since $Y|X \sim \mathcal{N}(x, \sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x))$ and $Z|X \sim \mathcal{N}(x, \sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x))$, (117) reduces to

$$\sqrt{\frac{\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)}{\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)}} \exp\left(\frac{(z-x)^2}{2\sigma_{\mathrm{E}}^2(1+\eta_{\mathrm{E}}^2 x)} - \frac{(y-x)^2}{2\sigma_{\mathrm{B}}^2(1+\eta_{\mathrm{B}}^2 x)} \right) = \frac{p(y|z)}{p(z|y)}. \tag{118}$$

To reach a contradiction, let us choose $y = z = 0$. For the contradiction, it is sufficient to show that the LHS of (118) is not a constant function in x. To this end, let us denote the LHS of (118) for $y = z = 0$ as $f(x)$. We show that $\frac{d[\log(f(x))]}{dx} < 0$ for all $x \in \mathcal{S}_{F_X}$.[2] The derivate of $\log(f(x))$ is given by

$$\begin{aligned} &\frac{d[\log(f(x))]}{dx} \\ &= \frac{\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2}{2}\left[\frac{(\sigma_{\mathrm{B}}^2 - \sigma_{\mathrm{E}}^2)}{(\sigma_{\mathrm{E}}^2 + \sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 x)(\sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2 x)}\right] \\ &\quad + x\left[\frac{(\sigma_{\mathrm{B}}^2 - \sigma_{\mathrm{E}}^2)}{(\sigma_{\mathrm{E}}^2 + \sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 x)(\sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2 x)}\right] \\ &\quad + \frac{x^2\,\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2}{2}\left[\frac{(\sigma_{\mathrm{E}}^4 - \sigma_{\mathrm{B}}^4) + 2\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2(\sigma_{\mathrm{E}}^2 - \sigma_{\mathrm{B}}^2)x}{(\sigma_{\mathrm{E}}^2 + \sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 x)^2(\sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2 x)^2}\right] \end{aligned}$$

[2] We note that for $y = z = 0$, $f(x) > 0$ for all $x \in \mathcal{S}_{F_X}$ and as a result, the sign of $\frac{df(x)}{dx}$ is the same as that of $\frac{d[\log(f(x))]}{dx}$.

$$= \frac{\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2}{2}\left[\frac{(\sigma_{\mathrm{B}}^2-\sigma_{\mathrm{E}}^2)}{(\sigma_{\mathrm{E}}^2+\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 x)(\sigma_{\mathrm{B}}^2+\sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2 x)}\right] + \frac{x^2\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2(\sigma_{\mathrm{B}}^4-\sigma_{\mathrm{E}}^4)+2x\sigma_{\mathrm{B}}^2\sigma_{\mathrm{E}}^2(\sigma_{\mathrm{B}}^2-\sigma_{\mathrm{E}}^2)}{2(\sigma_{\mathrm{E}}^2+\sigma_{\mathrm{E}}^2\eta_{\mathrm{E}}^2 x)(\sigma_{\mathrm{B}}^2+\sigma_{\mathrm{B}}^2\eta_{\mathrm{B}}^2 x)}. \quad (119)$$

Now, we note that since $\sigma_{\mathrm{E}}^2 > \sigma_{\mathrm{B}}^2$ and x is nonnegative (as x must satisfy the nonnegativity constraint), (119) is strictly negative for all $x \in \mathcal{S}_{F_X}$ and consequently, $\frac{df(x)}{dx} < 0$ for all $x \in \mathcal{S}_{F_X}$. This implies that for $y = z = 0$, $f(x)$ is not a constant function of x, which is a contradiction. This, in turn, implies that $I(Q; Y|Z) > 0$ and as a result, $I(X; Y|Z)$ is strictly concave in F_X. Furthermore, the output distributions are unique, i.e., if F_{X_1} and F_{X_2} are both secrecy-capacity-achieving, then $p_Y(y; F_{X_1}) = p_Y(y; F_{X_2})$ and $p_Z(z; F_{X_1}) = p_Z(z; F_{X_2})$.

Appendix B
Secrecy Capacity for Low-Intensity Regime Under Peak-Intensity Constraint

In the low-intensity regime, the secrecy capacity can be written as

$$C_{\mathrm{S}} = \frac{1}{2}[J_{\mathrm{B}}(0) - J_{\mathrm{E}}(0)] \max_{F_X \in \mathcal{A}^+} \mathrm{Var}(X) + o(A^2). \quad (120)$$

Therefore, the optimal input distribution that attains the secrecy capacity under nonnegativity and peak-intensity constraints in the low-intensity regime, also maximizes the variance of the input random variable $\mathrm{Var}(X)$. One can show that $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ is a strictly concave function in F_X. Since the set $\mathcal{A}^+$ is compact and convex and the functional $\mathrm{Var}(X)$ is continuous and strictly concave in F_X, the maximizer of the optimization problem in (120) exists and is unique. Moreover, the condition for the optimality of F_X^* is as follows [16, Ch. 12]

$$\frac{\delta\,\mathrm{Var}(X)}{\delta\, f_X^*(x)} = x^2 - 2\,x\int_0^A t\; f_X^*(t)\,dt = 0, \quad (121)$$

where $f_X^*(x)$ is the optimal probability density function (PDF) of random variable X and $\frac{\delta\,\mathrm{Var}(X)}{\delta\, f_X^*(x)}$ is the functional derivative of $\mathrm{Var}(X)$ with respect to $f_X^*(x)$. Now, let us assume that the optimal input distribution that satisfies (121) is $f_X^*(x) = p_0\,\delta(x - x_0) + p_1\,\delta(x - x_1)$, where $\delta(\cdot)$ is the dirac delta function. Substituting this distribution into (121) results in

$$x_0^2 - 2x_0^2\,p_0 - 2x_0x_1\,p_1 = 0, \quad (122)$$
$$x_1^2 - 2x_1^2\,p_1 - 2x_1x_0\,p_0 = 0. \quad (123)$$

One can verify that the optimal mass points are located at $\{x_0 = 0,\, x_1 = A\}$ and their corresponding probabilities are $\{p_0 = p_1 = 0.5\}$. Hence, $\max_{F_X \in \mathcal{A}^+} \mathrm{Var}(X) = \frac{A^2}{4}$.

Appendix C
Secrecy Capacity for Low-Intensity Regime Under Peak- and Average-Intensity Constraints

In the low-intensity regime, the secrecy capacity can be written as

$$C_{\mathrm{S}} = \frac{1}{2}[J_{\mathrm{B}}(0) - J_{\mathrm{E}}(0)] \max_{F_X \in \mathcal{M}^+} \mathrm{Var}(X) + o(A^2). \quad (124)$$

Therefore, the optimal input distribution that attains the secrecy capacity under both the peak- and average-intensity constraints for low-intensity regime, also maximizes the variance of the input random variable $\mathrm{Var}(X)$. Since the set $\mathcal{M}^+$ is compact and convex and the functional $\mathrm{Var}(X)$ is continuous and strictly concave in F_X, the maximizer of the optimization problem in (124) exists and is unique. Moreover, the optimization problem in (124) is equivalent to the following

$$h(f_X, \ell) = \max_{F_X \in \mathcal{M}^+} \mathrm{Var}(X) = \max_{F_X \in \mathcal{A}^+} \mathrm{Var}(X) - \ell(\mathbb{E}[X] - P), \quad (125)$$

where ℓ is the Lagrangian multiplier and positive. Therefore, the optimality conditions for F_X^* can be given by [16, Ch. 12]

$$\frac{\delta\, h(f_X^*, \ell)}{\delta\, f_X^*(x)} = 0, \quad (126)$$
$$\frac{\partial h(f_X^*, \ell)}{\partial \ell} = 0. \quad (127)$$

Next, let us consider that the optimal input distribution that satisfies (126)–(127) is $f_X^*(x) = p_0\,\delta(x - x_0) + p_1\,\delta(x - x_1)$, where $\delta(\cdot)$ is the dirac delta function. Substituting this distribution into (126)–(127) results in

$$x_0^2 - 2x_0^2\,p_0 - 2x_0x_1\,p_1 - \ell\,x_0 = 0, \quad (128)$$
$$x_1^2 - 2x_1^2\,p_1 - 2x_1x_0\,p_0 - \ell\,x_1 = 0, \quad (129)$$
$$x_0\,p_0 + x_1\,p_1 = P. \quad (130)$$

One can show that for $\ell = A(1 - 2\kappa)$ the optimal mass points are located at $\{x_0 = 0, x_1 = A\}$ and their corresponding probabilities are $\{p_0 = 1 - \kappa, p_1 = \kappa\}$. Thus, $\max_{F_X \in \mathcal{A}^+} \mathrm{Var}(X) = \kappa(1 - \kappa)A^2$. Since $\ell > 0$ (average intensity constraint is active), then $\kappa \in (0, \frac{1}{2})$. When $\kappa \in [\frac{1}{2}, 1]$ the average intensity constraint is not active and the result follows from Appendix B. That is, in the low-intensity regime, $\kappa \in (0, \frac{1}{2})$. A similar observation has been made in [3], but for the case with no secrecy constraint.

References

[1] S. Arnon, J. Barry, G. Karagiannidis, R. Schober, and M. Uysal, *Advanced Optical Wireless Communication Systems*, 1st ed. New York, NY, USA: Cambridge Univ. Press, 2012.

[2] A. Lapidoth, S. M. Moser, and M. A. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[3] S. M. Moser, "Capacity results of an optical intensity channel with input-dependent Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.

[4] 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.

[5] C. E. Shannon, "Communication theory of secrecy systems," *Bell Labs Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

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

[7] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.

[8] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[9] O. Ozel, E. Ekrem, and S. Ulukus, "Gaussian wiretap channel with amplitude and variance constraints," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5553–5563, Oct. 2015.

[10] J. G. Smith, "The information capacity of amplitude- and variance-constrained sclar Gaussian channels," *Inf. Control*, vol. 18, no. 3, pp. 203–219, 1971.
[11] Z. Wang, Q. Wang, W. Huang, and Z. Xu, *Visible Light Communications: Modulation and Signal Processing*, 1st ed. Piscataway, NJ, USA: Wiley, Nov. 2017.
[12] C. Luo, "Communication for wideband fading channels: On theory and practice," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol.,Cambridge, MA, USA, Feb. 2006.
[13] R. M. Dudley, *Real Analysis and Probability*, vol. 74. Cambridge, U.K.: Cambridge Univ. Press, 2002.
[14] V. V. Prelov and E. C. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1728–1735, Sep. 1993.
[15] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.
[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.
[17] K. A. Dadamahalleh and G. A. Hodtani, "A general upper bound for FSO channel capacity with input-dependent Gaussian noise and the corresponding optimal input distribution," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1700–1704.

Morteza Soltani (S'18) was born in Mashhad, Iran in 1991. He received his BS.c. degree from the Ferdowsi University of Mashhad, Mashhad, Iran, in 2013 and his MS.c. degree from the Istanbul Medipol University, Istanbul, Turkey, in 2016. He is now pursuing a Ph.D. degree in electrical engineering at the University of Idaho, Moscow, Idaho, USA. Morteza's research interests are Information-Theoretic Security, Network Information Theory and Optical Communications.

Zouheir Rezki (S'01–M'08–SM'13) was born in Casablanca, Morocco. He received the Diplôme d'Ingénieur degree from the École Nationale de l'Industrie Minérale (ENIM), Rabat, Morocco, in 1994, the M.Eng. degree from École de Technologie Supérieure, Montreal, Québec, Canada, in 2003, and the Ph.D. degree in electrical engineering from École Polytechnique, Montreal, Québec, in 2008. After a few years of experience as a postdoctoral fellow and a research scientist at KAUST, he joined the University of Idaho as an Assistant Professor in the ECE Department.