The Discrete-Time Poisson Optical Wiretap Channel with Peak Intensity Constraint

Morteza Soltani, Student Member, IEEE, and Zouheir Rezki, Senior Member, IEEE
University of Idaho, Department of Electrical and Computer Engineering
83844 Moscow, Idaho
Emails: {msoltani, zrezki}@uidaho.edu

Abstract

This paper addresses the discrete-time Poisson wiretap channel (DT–PWC) in an optical wireless communications system based on intensity modulation and direct detection. Subject to nonnegativity and peak intensity as well as bandwidth constraints imposed on the channel input, we study the secrecy-capacity-achieving input distribution of this wiretap channel and prove it to be unique and discrete with a finite number of mass points. Furthermore, we establish that every point on the boundary of the rate-equivocation region of this wiretap channel is also obtained by a unique and discrete input distribution with a finite support. In general, the number of mass point of the optimal distributions are greater than two. This is in contrast with the continuous-time PWC where the secrecy capacity and the entire boundary of the rate-equivocation region are achieved by binary distributions when the signaling bandwidth is not restricted. Additionally, 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. Our numerical results indicate that there is a tradeoff between the secrecy capacity and the capacity in the sense that both may not be achieved simultaneously.

I. INTRODUCTION

Optical wireless communications is often based on intensity modulation and direct detection (IM-DD), where 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 which absorbs integer number of photons and generates a real valued output
corrupted by noise. Based on the distribution of this corrupting noise, there are several channel models for the underlying optical wireless channels. Free space optical (FSO) channels [1], [2], optical channels with input-dependent Gaussian noise [2], [3] and Poisson optical channels [2], [4]–[6] are the most widely used channel models for optical wireless communications. Among the existing channel models for optical wireless communications, the most accurate model that can capture most of the optical channel impairments, is the Poisson channel model. Considering a discrete-time Poisson channel, Shamai [5] studies the single-user channel capacity of this model and shows that the capacity-achieving distribution, under nonnegativity, peak and average intensity constraints, is discrete with a finite number of mass points. In [6], [7] authors provide asymptotic analysis of the channel capacity for the single-user case in the low and high intensity regime.

The broadcast nature of optical wireless signals imposes a security challenge especially in the presence of unauthorized eavesdroppers. This problem has been conventionally addressed by cryptographic encryption [8] without considering the imperfections introduced by the communication channel. Wyner [9], on the other hand, proved the possibility of secure communications without relying on encryption by introducing the notion of a degraded wiretap channel. These result has been later generalized by Csiszar and Korner by dropping the degradedness assumption of the wiretap channel [10].

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 at a certain secrecy level against an eavesdropper [11]. For the class of degraded wiretap channels, it is established in [9] that there exists a single-letter characterization for the rate-equivocation region. Authors in [12] study the degraded FSO wiretap channel under a peak intensity constraint and prove that the entire rate-equivocation region of this wiretap channel is achieved by discrete input distributions with a finite number of mass points. Furthermore, the authors observe that 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. The work in [13] considers the degraded optical wiretap channel with input-dependent Gaussian noise under peak and average intensity constraints and establishes the optimality of discrete input distributions with a finite support for attaining the entire boundary of the rate-equivocation region. Moreover, the authors provide asymptotic behavior of the secrecy capacity in the high and low power regimes. For this wiretap channel, authors similarly observe
that, in general, there is a tradeoff between the rate and its equivocation. Finally, [14] examines the continuous-time degraded Poisson wiretap channel with a peak intensity constraint, and gives a closed-form expression for the secrecy-capacity. Particularly, the authors show that a binary input distribution with a very short duty cycle can achieve the secrecy capacity of the continuous-time degraded wiretap channel.

In this work, we consider a discrete-time degraded Poisson wiretap channel (DT–PWC) which consists of a transmitter, a legitimate user and an eavesdropper. In this wiretap channel, the input signals are restricted to have finite bandwidth. This fact distinguishes the DT–PWC from its continuous-time counterpart, where input signals can have infinite bandwidth. Using an IM-DD system, the photodetectors at the legitimate user and the eavesdropper counts the number of received photons and output signals that follow Poisson distribution. In this setup, the objective is to have secure communication with the legitimate user over an optical channel while keeping the eavesdropper ignorant of the transmitted message as much as possible. To this end, we start by the secrecy capacity of the degraded DT–PWC and employ the functional optimization problem addressed in, for example [5], [12], [13], [15], to obtain the necessary and sufficient conditions, also known as Karush-Kuhn-Tucker (KKT) conditions, for the optimal input distribution. We then resort to a contradiction argument to prove that the secrecy capacity is achieved by a unique discrete input distribution with a finite number of mass points. Moreover, we extend the optimality of the discrete input distributions with a finite number of mass points to the entire boundary of the rate-equivocation region of the DT–PWC. Additionally, we provide an asymptotic analysis for the secrecy capacity in the low and high intensity regimes. We observe that the in the low intensity regime, the secrecy capacity scales quadratically in the peak intensity constraint. In the high-intensity regime, on the contrary, the secrecy capacity can be upper-bounded by a positive constant implying that it does not scale with the constraint. Through our numerical experiments, we observe that similar to the cases of the FSO wiretap channel and the optical intensity wiretap channel with input-dependent noise with a peak intensity constraint, here too, the secrecy capacity and the capacity are not achieved by the same distributions in general. This, in turn, implies that for this wiretap channel, there is a tradeoff between the rate and its equivocation.

The rest of the paper is organized as follows. The discrete-time degraded Poisson wiretap channel is formally defined in Section II. The main result of the work regarding the characterization of the optimal input distributions that achieve the secrecy capacity as well as the entire
boundary of the rate-equivocation region is presented in Section III. Proof of the main result is provided in Section IV. Asymptotic analysis of the secrecy capacity in the low and high intensity regimes are demonstrated in Section V. Numerical results are shown in Section VI, and finally, conclusions are drawn in Section VII.

II. DISCRETE-TIME POISSON WIRETAP CHANNEL

In the considered optical wiretap channel, the channel input $X$ is a nonnegative random variable representing the intensity of the optical signal. Since intensity is constrained due to practical and safety restrictions by a peak constraint in general, the input has to satisfy [2]

$$0 \leq X \leq A.$$  \hfill (1)

In this setup, conditional on the input $x$, the output signal at the legitimate receiver $Y$ and at the eavesdropper $Z$ is Poisson distributed with mean $x + \lambda_B$ and $x + \lambda_E$, respectively, where $\lambda_B$ and $\lambda_E$ are some nonnegative constants, called dark currents. Thus, the conditional channel laws are given by [6]

$$p_{Y|X}(y|x) = e^{-(x+\lambda_B)} \frac{(x+\lambda_B)^y}{y!}, \quad y \in \mathbb{N},$$  \hfill (2)

$$p_{Z|X}(z|x) = e^{-(x+\lambda_E)} \frac{(x+\lambda_E)^z}{z!}, \quad z \in \mathbb{N},$$  \hfill (3)

where $\mathbb{N}$ is the set of nonnegative integers. In this wiretap channel, if the dark current in the eavesdropper’s channel is greater than that in the legitimate user’s channel, i.e., $\lambda_E > \lambda_B$, the random variables $X$, $Y$, and $Z$ form the Markov chain $X \rightarrow Y \rightarrow Z$ and consequently the DT–PWC becomes stochastically degraded [9], [14], [16]. Otherwise, the secrecy capacity, defined later in this section, is zero.

An $(n,2^{nR})$ code for the DT–PWC with peak intensity constraint consists of the random variable $W$ (message set) uniformly distributed over the set $W = \{1,2,\cdots,2^{nR}\}$, an encoder at the transmitter $f_n : W \rightarrow [0,A]^n$ satisfying the nonnegativity and peak intensity constraints, and a decoder at the legitimate user $g_n : \mathbb{N}^n \rightarrow W$. Equivocation of a code 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 $P_e^n = \Pr[g_n(Y^n) \neq W]$. A rate-equivocation pair $(R,R_e)$ is said to be achievable if there exists an $(n,2^{nR})$ code satisfying

$$\lim_{n \to \infty} P_e^n = 0,$$  \hfill (4)

$$R_e \leq \lim_{n \to \infty} \frac{1}{n} H(W|Z^n),$$  \hfill (5)
where \( H(W|Z^n) \) is the conditional entropy of \( W \) given the observations \( Z^n \). The rate-equivocation region consists of all achievable rate-equivocation pairs. A rate \( R \) is said to be perfectly secure if we have \( R_e = R \), i.e., if there exists an \((n, 2^{nR})\) code satisfying \( \lim_{n \to \infty} \frac{1}{n} I(W; Z^n) = 0 \), where \( I(W; Z^n) \) is the mutual information between the random variables \( W \) and \( Z^n \). The supremum of such rates is defined to be the secrecy capacity denoted by \( C_S \).

Since under the assumption of \( \lambda_E > \lambda_B \), the DT–PWC with nonnegativity and peak intensity constraints is degraded, its entire rate-equivocation region, denoted by \( \mathcal{R} \), can be expressed in a single-letter expression and the entire rate-equivocation region of this wiretap channel is given by the union of the rate-equivocation pairs \((R, R_e)\) such that [9]

\[
\begin{align*}
0 \leq R &\leq I(X; Y), \\
0 \leq R_e &\leq I(X; Y) - I(X; Z).
\end{align*}
\] (6)

for some input distribution \( F_X \in \mathcal{A}^+ \), where the feasible set \( \mathcal{A}^+ \) is given by

\[
\mathcal{A}^+ \triangleq \left\{ F_X : \int_0^A dF_X(x) = 1 \right\}.
\] (7)

III. MAIN RESULTS

In this section, we present the main results related to the DT–PWC. 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 a finite support.

A. Secrecy Capacity

For the degraded DT–PWC, the secrecy capacity is given by a single-letter expression as [11, Chap. 3]

\[
C_S = \sup_{F_X \in \mathcal{A}^+} f_0(F_X) \triangleq \sup_{F_X \in \mathcal{A}^+} [I(X; Y) - I(X; Z)].
\] (8)

The secrecy-capacity-achieving of the DT–PWC with nonnegativity and peak intensity constraints is given by the solution of the optimization problem in (8). Under the constraint (1), the solution to (8) exists, is unique and discrete with a finite support. This is formally presented by Theorem 1.

**Theorem 1.** There exists a unique input distribution that attains the secrecy capacity of the DT–PWC with nonnegativity and peak intensity constraints. Furthermore, the support set of this optimal input distribution is a finite set.
Proof. For convenience, the proof is presented in Section IV.

It is worth mentioning that in the continuous-time Poisson wiretap channel studied in [14], the secrecy-capacity-achieving input distribution is always discrete with two mass points located at the origin and the value of the peak intensity constraint. Furthermore, to achieve the secrecy capacity, input signals must have very short duty cycle (or equivalently, a very large transmission bandwidth is required). However, in the DT-PWC, the number of mass points of the secrecy-capacity-achieving input distribution depends on the value of the peak intensity constraint and in general, it is larger than two. Moreover, the input signals are restricted with respect to the signalling bandwidth.

B. Rate-Equivocation Region

By a time-sharing argument, it can be shown that the rate-equivocation region of the DT–PWC is convex. Therefore, the region can be characterized by finding tangent lines to $\mathcal{R}$, which are given by the solutions of

$$
\max_{F_X \in \mathcal{A}^+} f_0(F_X) = \alpha I(X;Y) + (1 - \alpha) [I(X;Y) - I(X;Z)], \quad \forall \alpha \in [0,1].
$$

Next, we establish that the entire rate-equivocation region of the DT–PWC with the constraint (1) is also obtained by discrete input distributions with a finite number of mass points.

Theorem 2. Every point on the boundary of the rate-equivocation region of the DT-PWC with nonnegativity and peak intensity constraints, is achieved by a unique input distribution which is discrete with a finite support.

Proof. For brevity, Theorem 2 is established in Section IV.

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 Theorem 1 and Theorem 2. We start by proving that the set of input distributions $\mathcal{A}^+$ satisfying (7), is compact and convex. We then show that the objective functions $f_0(F_X)$ and $f_\alpha(F_X)$ in (8) and (9), respectively, are continuous, strictly concave and weakly differentiable in the input distribution $F_X$ and hence we conclude that the optimization problems in (8) and (9) have unique solutions. We continue the proofs by deriving the necessary and sufficient conditions (KKT conditions) for the optimality of the optimal input
distribution \( F_X \). Finally, by means of contradiction we show that the optimal input distributions are discrete with a finite number of mass points.

A. Preliminaries

Since both channels are discrete-time Poisson, the output densities for \( Y \) and \( Z \) exist for any input distribution \( F_X \), and are given by

\[
P_Y(y; F_X) = \int_{0}^{\Lambda} p(y|x) \, dF_X(x), \quad y \in \mathbb{N} \tag{10}
\]

\[
P_Z(z; F_X) = \int_{0}^{\Lambda} p(z|x) \, dF_X(x), \quad z \in \mathbb{N} \tag{11}
\]

where \( p(y|x) \) and \( p(z|x) \) are given by (2)–(3).

We define the secrecy rate density \( r_e(x; F_X) \) as

\[
r_e(x; F_X) = i_B(x; F_X) - i_E(x; F_X), \tag{12}
\]

where \( i_B(x; F_X) \) and \( i_E(x; F_X) \) are the mutual information densities for the legitimate user’s and the eavesdropper’s channel, respectively, and are given by

\[
i_B(x; F_X) = \sum_{y=0}^{+\infty} p(y|x) \log \left( \frac{p(y|x)}{P_Y(y; F_X)} \right), \tag{13}
\]

\[
i_E(x; F_X) = \sum_{z=0}^{+\infty} p(z|x) \log \left( \frac{p(z|x)}{P_Z(z; F_X)} \right). \tag{14}
\]

Plugging (2)–(3) and (10)–(11) into (13)–(14) and after some algebra, we get

\[
i_B(x; F_X) = (x + \lambda_B) \log(x + \lambda_B) - x - \sum_{y=0}^{+\infty} p(y|x) \log(g_B(y; F_X)), \tag{15}
\]

\[
i_E(x; F_X) = (x + \lambda_E) \log(x + \lambda_E) - x - \sum_{z=0}^{+\infty} p(z|x) \log(g_E(z; F_X)), \tag{16}
\]

where \( g_B(y; F_X) \) and \( g_E(z; F_X) \) are respectively defined as

\[
g_B(y; F_X) \triangleq \int_{0}^{\Lambda} e^{-x} (x + \lambda_B)^y \, dF_X(x), \tag{17}
\]

\[
g_E(z; F_X) \triangleq \int_{0}^{\Lambda} e^{-x} (x + \lambda_E)^z \, dF_X(x). \tag{18}
\]
Furthermore, we have the following identities

\begin{align}
I(X; Y) &= \int_{0}^{A} i_B(x; F_X) \, dF_X(x) \overset{\Delta}{=} I_B(F_X), \\
I(X; Z) &= \int_{0}^{A} i_E(x; F_X) \, dF_X(x) \overset{\Delta}{=} I_E(F_X), \\
f_0(F_X) &= \int_{0}^{A} r_e(x; F_X) \, dF_X(x).
\end{align}

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

**B. Proof of Theorem 1**

First, for convenience, let us designate by $I_B(F_X)$ and $I_E(F_X)$ the mutual informations $I(X; Y)$ and $I(X; Z)$, respectively, induced by the input distribution $F_X$. The proof is then streamlined into a few lemmas which we state below.

**Lemma 1.** The feasible set $\mathcal{A}^+$ is convex and sequentially compact in the Levy metric sense.

*Proof.* The proof follows along similar lines as [5, Lemma 1].

**Lemma 2.** The functional $f_0 : \mathcal{A}^+ \to \mathbb{R}$, $f_0(F_X) = I_B(F_X) - I_E(F_X)$ is continuous in $F_X$.

*Proof.* The proof follows along similar lines as presented in [5, Lemma 3].

From Lemma 1 and Lemma 2, $f_0(F_X)$ is continuous in $F_X$ over $\mathcal{A}^+$ which itself is a compact set, then by the Extreme Value Theorem, $f_0(F_X)$ is bounded above and attains its supremum. That is, the supremum in (8) is actually a maximum which is achievable by at least one input distribution $F_X$.

**Lemma 3.** The functional $f_0(F_X)$ is strictly concave in $F_X$.

*Proof.* The proof is by contradiction and follows along similar lines as in [13, Appendix A] with the difference that the conditional channel laws follow Poisson distribution. For completeness, the proof is relegated to Appendix A.

**Lemma 4.** The functional $f_0(F_X)$ is weakly differentiable in $\mathcal{A}^+$ and its weak derivative at the point $F_X^o$, denoted by $f'_0(F_X^o)$ is given by

\[ f'_0(F_X^o) = \int_{0}^{A} r_e(x; F_X^o) \, dF_X - f(F_X^o). \]
Proof. The proof is based on the definition of the weak derivative and follows along similar lines as the one in [13].

From Lemma 1, Lemma 3, and Lemma 4, we have a strictly concave and weak-differentiable function \( f_0(F_X) \) over \( \mathcal{A}^+ \) which is a convex set, then a sufficient and necessary condition for an input distribution \( F_X^* \) to be optimal is

\[
f_0'(F_X^*) \leq 0,
\]

for any \( F_X \in \mathcal{A}^+ \). Steps analogous to [5, Lemma 4] or [15, Corollary 1] yield the following necessary and sufficient conditions for the optimality of \( F_X^* \)

\[
r_e(x; F_X^*) \leq C_S, \quad \forall x \in [0, A] \tag{24}
\]

\[
r_e(x; F_X^*) = C_S, \quad \forall x \in S_{F_X^*} \tag{25}
\]

where \( S_{F_X^*} \subset [0, A] \) is the support set of \( F_X^* \) and the secrecy capacity is \( C_S = I_B(F_X^*) - I_E(F_X^*) = f_0(F_X^*) \).

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 (24)–(25). To this end, the following lemma establishes that both \( i_B(w; F_X) \) and \( i_E(w; F_X) \) have analytic extensions over some open connected set in the complex plane \( \mathbb{C} \).

Lemma 5. The secrecy rate density \( r_e(x; F_X) \) has an analytic extension to the open connected set \( O \triangleq \{ w \in \mathbb{C} : \Re(w) > -\lambda_B \} \), where \( \Re(w) \) is the real part of the complex variable \( w \).

Proof. The mutual information densities \( i_B(w; F_X) \) and \( i_E(w; F_X) \) have analytic extension to the open connected sets \( O_B \triangleq \{ w \in \mathbb{C} : \Re(w) > -\lambda_B \} \) and \( O_E \triangleq \{ w \in \mathbb{C} : \Re(w) > -\lambda_E \} \), respectively, according to [5]. Therefore, the secrecy rate density \( r_e(w; F_X) \) has an analytic extension to the open connected set \( O = O_B \cap O_E \). Since \( \lambda_E > \lambda_B \), we get \( O = O_B \). This completes the proof of Lemma 5.

Now, we are ready to prove the discreteness and finiteness of the support set of \( F_X^* \) using a contradiction argument. We start by assuming that \( S_{F_X^*} \) has an infinite number of elements. In view of the optimality condition (25), the analyticity of \( r_e(w; F_X) \) over \( O \) and the Identity Theorem from complex analysis along with Bolzano-Weierstrass Theorem, if \( S_{F_X^*} \) has an infinite
number of mass points, we deduce that \( r_e(w; F^+_X) = C_S \) for all \( w \in O \). Since \((-\lambda_B, +\infty) \subset O\), we conclude that

\[
r_e(x; F^+_X) = C_S, \quad \forall x \in (-\lambda_B, +\infty).
\]

(26)

Next, we show that (26) results in a contradiction. Observe that (26) implies that \( r_e(x; F^+_X) \) is a constant function in \( x \) for all \( x \in (-\lambda_B, +\infty) \). Therefore, to reach a contradiction, we show that \( r_e(x; F^+_X) \) is not a constant function over the interval \((-\lambda_B, +\infty)\). To this end, we take the derivative of the sides of (26) with respect to \( x \) and we find

\[
\frac{dr_e(x; F^+_X)}{dx} = 0, \quad \forall x \in (-\lambda_B, +\infty).
\]

(27)

Substituting (15)–(16) into (12) and taking the derivative with respect to \( x \), we can write

\[
\frac{dr_e(x; F^+_X)}{dx} = \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] + e^{-(x+\lambda_B)} \log (g_B(0; F^+_X)) + e^{-(x+\lambda_E)} \log (g_E(0; F^+_X)) + \sum_{y=1}^{+\infty} |p(y|x) - \rho(y-1|x)\log (g_B(y; F^+_X)) + \sum_{z=1}^{+\infty} -\rho(z-1|x)\log (g_E(z; F^+_X)).
\]

(28)

It can be easily shown that \( g_B(y; F^+_X) \) and \( g_E(z; F^+_X) \) are bounded as follows:

\[
e^{-A} \lambda_B^y \leq g_B(y; F^+_X) \leq (A + \lambda_B)^y,
\]

(29)

\[
e^{-A} \lambda_E^z \leq g_E(z; F^+_X) \leq (A + \lambda_E)^z.
\]

(30)

Using the bounds in (29)–(30), \( \frac{dr_e(x; F^+_X)}{dx} \) can be upper bounded by

\[
\frac{dr_e(x; F^+_X)}{dx} \leq \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] + (x + \lambda_B) \log \left[ 1 + \frac{A}{\lambda_B} \right] + (x + \lambda_E) \log \left[ 1 + \frac{A}{\lambda_E} \right] + \log \left[ \frac{e^A (A + \lambda_E)}{\lambda_B} \right], \quad \forall x \in (-\lambda_B, +\infty).
\]

(31)

Finally, to reach a contradiction, it suffices to compute the limit of the right hand side of (31) as \( x \to -\lambda_B^+ \). For this purpose, and in regard of (27), we have

\[
0 \leq \lim_{x \to -\lambda_B^+} \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] + 0 + (\lambda_E - \lambda_B) \log \left[ 1 + \frac{A}{\lambda_E} \right] + \log \left[ \frac{e^A (A + \lambda_E)}{\lambda_B} \right], \quad \forall x \in (-\lambda_B, +\infty).
\]

(32)

We observe that since \( \lambda_E > \lambda_B \), the limit \( \lim_{x \to -\lambda_B^+} \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] = -\infty \) and therefore, we get 0 \( \leq -\infty \) which is a contradiction. Therefore, \( r_e(x; F^+_X) \) cannot be a constant function over \((-\lambda_B, +\infty)\).
This, in turn, implies that the support set \( S_{F_X^*} \) 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 Theorem 2

This section presents the proof of Theorem 2 by extending the analysis in the previous section to the entire rate-equivocation region. We start by noting that the objective function \( f_o(F_X) \) in (9) is strictly concave, and the feasible set \( \mathcal{A}^+ \) is compact and convex, therefore, the optimization problem in (9) has a unique maximizer. We denote the optimal input distribution for (9) by \( F_X^* \) which depends on the value \( \alpha \).

Now, we obtain the KKT conditions for the optimal input distribution of the optimization problem in (9). We note that the objective function \( f_o(F_X) \) is weakly differentiable and its weak derivative is given by

\[
f'_o(F_X^*) = \int_0^\Lambda [\alpha i_B(x; F_X^*) + (1 - \alpha) r_e(x; F_X^*)] dF_X(x) - f_o(F_X^*). \tag{33}
\]

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

\[
\alpha i_B(x; F_X^*) + (1 - \alpha) r_e(x; F_X^*) \leq \alpha I_B(F_X^*) + (1 - \alpha) \left[ I_B(F_X^*) - I_E(F_X^*) \right], \quad \forall x \in [0, \Lambda], \tag{34}
\]

\[
\alpha i_B(x; F_X^*) + (1 - \alpha) r_e(x; F_X^*) = \alpha I_B(F_X^*) + (1 - \alpha) \left[ I_B(F_X^*) - I_E(F_X^*) \right], \quad \forall x \in S_{F_X^*}. \tag{35}
\]

Next, we show that the optimal input distribution \( F_X^* \) has a finite support. To this end, assume to the contrary, that \( S_{F_X^*} \) has an infinite number of elements. Under such an assumption, (35), the analyticity of \( i_B(w; F_X^*) \) and \( i_E(w; F_X^*) \) over \( \mathcal{O} \) in the complex plane and the Identity Theorem of complex analysis imply that

\[
\alpha i_B(x; F_X^*) + (1 - \alpha) r_e(x; F_X^*) = \alpha I_B(F_X^*) + (1 - \alpha) \left[ I_B(F_X^*) - I_E(F_X^*) \right], \quad \forall x \in (-\lambda_B, +\infty). \tag{36}
\]

We continue the proof by showing that (36) results in a contradiction. To do so, we first observe that the right hand side (RHS) of (36) does not depend on \( x \), and hence, it is a constant function in \( x \). Let us designate the RHS of (36) by \( c \). Plugging (15)–(16) into (36) and after some simplifications, we get

\[
\alpha(x + \lambda_E) \log(x + \lambda_E) - \alpha x + (x + \lambda_B) \log(x + \lambda_B) - (x + \lambda_E) \log(x + \lambda_E)
\]

\[
= \sum_{y=0}^{+\infty} p(y|x) \log(g_B(y; F_X^*)) - (1 - \alpha) \sum_{z=0}^{+\infty} p(z|x) \log(g_E(y; F_X^*)) + c. \tag{37}
\]
Using the bounds in (29)–(30), the RHS of (37) can be upper bounded as
\[
\sum_{y=0}^{+\infty} p(y|x) \log(g_B(y; F_X^*)) - (1 - \alpha) \sum_{z=0}^{+\infty} p(z|x) \log(g_E(y; F_X^*)) + c \leq (x + \lambda_B) \log(x + \lambda_B) \\
- (1 - \alpha)(x + \lambda_E) \log(\lambda_E) + (1 - \alpha)A + c.
\] (38)

Combining (37) with (38) and after some simplification, we get
\[
\alpha(x + \lambda_E) \log(x + \lambda_E) + (x + \lambda_B) \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] + (\lambda_B - \lambda_E) \log(x + \lambda_E) \\
\leq x \left[ \alpha + \log(A + \lambda_B) - (1 - \alpha) \log(\lambda_E) \right] + \lambda_B \log(A + \lambda_B) - (1 - \alpha)\lambda_E \log(\lambda_E) + (1 - \alpha)A + c, \\
\forall x \in (-\lambda_B, +\infty). 
\] (39)

Now, observe that the RHS of (39) cannot grow faster than \(x \left[ \alpha + \log(A + \lambda_B) - (1 - \alpha) \log(\lambda_E) \right] \) for \(\alpha \in (0, 1)\). On the other hand, the left hand side of (39) grows as \(x \log x\), which leads to a contradiction. This implies that \(S_{F_X^*}\) cannot have infinite elements. Therefore, \(F_X^*\) is discrete with a finite number of mass points. Finally, we note that for \(\alpha = 0\), \(F_X^*\) must be discrete with a finite support according to Theorem 1 and for \(\alpha = 1\), \(F_X^*\) is also discrete with a finite number of mass points based on [5]. Consequently, the entire boundary of the rate-equivocation region of the DT–PWC is obtained by discrete input distributions with a finite support. This completes the proof of Theorem 2.

V. ASYMPTOTIC RESULTS FOR A PEAK-INTENSITY CONSTRAINT

This section provides the asymptotic analysis on the secrecy capacity of the DT–PWC with 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 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 [17] and we can write
\[
I(X; Y) - I(X; Z) = \frac{1}{2} [J_B(0) - J_E(0)] \text{Var}(X) + o(A^2),
\] (40)
where \( o(A^2) \) denotes a function satisfying \( \lim_{A \to 0} \frac{o(A^2)}{A^2} = 0 \), \( \text{Var}(X) \) is the variance of the random variable \( X \), \( J_B(0) \) and \( J_E(0) \) denote the Fisher information of the legitimate user’s and the eavesdropper’s channel at 0 with \( J_B(x) \) and \( J_E(x) \) given by

\[
J_B(x) = \frac{1}{x + \lambda_B}, \quad (41)
\]
\[
J_E(x) = \frac{1}{x + \lambda_E}. \quad (42)
\]

As a result, we have \( J_B(0) = \frac{1}{\lambda_B} \) and \( J_E(0) = \frac{1}{\lambda_E} \). Consequently, for small values of \( A \), the secrecy capacity is

\[
C_S = \frac{1}{2} [J_B(0) - J_E(0)] \max_{F_X \in \mathcal{A}^+} \text{Var}(X) + o(A^2). \quad (43)
\]

**Proposition 1.** In the regime \( A \to 0 \), the secrecy capacity under the peak intensity constraint is as follows

\[
C_S(A) = \frac{A^2}{8} \left[ \frac{1}{\lambda_B} - \frac{1}{\lambda_E} \right] + o(A^2). \quad (44)
\]

Furthermore, under the constraints (1), two mass points located at 0 and \( A \) with equal probabilities are optimal in this regime.

**Proof.** See Appendix B.

Proposition 1 indicates that the secrecy capacity is a quadratic function of the peak intensity constraint \( A \) in the low intensity regime.

**B. High-Intensity Results**

In this section, we provide an upper bound on the secrecy capacity that holds for any value of \( A \) and any input distribution \( F_X \in \mathcal{A}^+ \). Based on the definition of the secrecy capacity in (8), the secrecy capacity can be written as

\[
C_S = H_Y(F_X^*) - H_Z(F_X^*) + H_{Z|X}(F_X^*) - H_{Y|X}(F_X^*), \quad (45)
\]

where \( H_Y(F_X^*) \) and \( H_Z(F_X^*) \) are the entropies of the random variables \( Y \) and \( Z \), respectively, induced by the optimal input distribution \( F_X^* \). Moreover, \( H_{Y|X}(F_X^*) \) and \( H_{Z|X}(F_X^*) \) are the conditional entropies of \( Y|X \) and \( Z|X \), respectively, induced by \( F_X^* \). Now, the following proposition provides an upper bound on the secrecy capacity that holds true for any value of \( A \) including the case where \( A \to \infty \).
Proposition 2. The secrecy capacity of the DT–PWC with nonnegativity and peak intensity constraint can be upper bounded by

\[ C_S \leq \frac{\lambda_D^2}{2} + \lambda_D, \tag{46} \]

where \( \lambda_D = \Lambda E - \Lambda B \).

Proof. For convenience, the proof of Proposition 2 is relegated to Appendix C. ■

Proposition 2 implies that the secrecy capacity in the regime \( A \to \infty \) does not scale with the peak intensity constraint and converges to a real and positive constant, i.e.,

\[ C_S(A) = O(1). \tag{47} \]

VI. Numerical Results

In this section, we provide numerical results for the secrecy capacity and the entire rate-equivocation region of the DT–PWC with nonnegativity and peak intensity constraints.

Figure 1 provides a plot of the KKT conditions given by (24)–(25) for an optimal input distribution when \( A = 5, \lambda_B = 1, \lambda_E = 2 \). We numerically found that for these parameters, the optimal input distribution is ternary with mass points located at \( x = 0, 2.8065 \) and \( 5 \) with probability masses \( 0.4847, 0.3014 \) and \( 0.2140 \), respectively. We observe that \( C_S - r_e(x; F^*_X) \) is generally nonnegative and is equal to zero at the optimal mass points; verifying the optimality conditions in (24)–(25).

Figure 2 illustrates the secrecy capacity \( C_S \) and the difference \( C_B - C_E \) versus the peak intensity constraint \( A \), where \( C_B \) and \( C_E \) are the legitimate user’s and the eavesdropper’s capacities, respectively. We observe that this difference is a lower bound on the secrecy capacity \( C_S \). We also observe that, for small values of \( A \), \( C_B - C_E \) and \( C_S \) are identical. However, as \( A \) increases, \( C_B - C_E \) and \( C_S \) become different. Similar to the secrecy capacity results of the FSO wiretap channel and optical wiretap channel with input-dependent Gaussian noise under a peak intensity constraint provided in [12], [13], 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 \( A = 5 \), while both \( I(X;Y) \) and \( I(X;Z) \) are maximized by the same binary distribution with mass points at \( x = 0 \) and \( 5 \) with probability masses \( 0.5118 \) and \( 0.4882 \), respectively, \( I(X;Y) - I(X;Z) \) is maximized by a ternary distribution with mass points at \( x = 0, 2.8065, \) and \( 5 \) with probability masses \( 0.4847, 0.3014, \) and \( 0.2140 \), respectively.
Fig. 1. Illustration of $C_S - r_x(x, F_X^*)$ yielded by the optimal input distribution when $\lambda_B = 1$, $\lambda_E = 2$, and $A = 5$.

Fig. 2. The secrecy capacity for $\lambda_B = 1$, $\lambda_E = 2$ versus the peak intensity constraint $A$.

0, 2.8065 and 5 with probability masses 0.4847, 0.3014 and 0.2140, respectively. This explains the difference between $C_S$ and $C_B - C_E$ at $A = 5$ in this figure.

Figure 3 depicts the entire rate-equivocation region of the DT–PWC with nonnegativity and peak intensity constraints when $\lambda_B = 1$, $\lambda_E = 2$, for two different values of $A$. When $A = 3.5$, 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 $A = 3.5$, the binary input distribution with mass points located at $x = 0$ and 3.5 with probabilities 0.5174 and 0.4826, respectively, achieves
both the capacity and the secrecy capacity. This implies that, when $A = 3.5$, the transmitter can communicate with the legitimate user at the capacity while achieving the maximum equivocation at the eavesdropper. On the other hand, when $A = 5$, the secrecy capacity and the capacity cannot be achieved simultaneously (notice the curved shape in the figure). More specifically, for $A = 5$, the binary input distribution with mass points located at $x = 0$ and $5$ with probabilities $0.5113$ and $0.4877$ achieves the capacity, while a ternary distribution with mass points located at at $x = 0$, $2.8065$ and $5$ with probability masses $0.4847$, $0.3014$ and $0.2140$, 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 on the equivocation of this communication, and to increase the achieved equivocation, one must compromise on the communication rate.

Finally, in Fig. 4, we plot the exact and asymptotic secrecy capacity results versus the peak intensity constraint $A$ for $\lambda_B = 1, \lambda_E = 1.25$ and $\lambda_B = 1, \lambda_E = 2$ in the low-intensity regime. From the figure, we observe that our asymptotic results for the secrecy capacity given in (44) are in precise agreement with the numerical results. Furthermore, we observe that increasing the dark current in the eavesdropper’s channel results in a higher secrecy capacity.
Fig. 4. The asymptotic and exact secrecy capacity for $\lambda_B = 1, \lambda_E = 2$ and $\lambda_B = 1, \lambda_E = 1.25$ versus the peak intensity constraint $A$.

VII. CONCLUSIONS

We studied the DT–PWC with nonnegativity and peak intensity constraints. We formally characterized the secrecy-capacity-achieving input distribution to be unique and discrete with a finite number of mass points. Furthermore, we established that every point on the boundary of the rate-equivocation region of this wiretap channel is also obtained by a unique and discrete input distribution with a finite number of mass points. The DT–PWC is in contrast to its continuous counterpart in the sense that the number of mass point of the optimal distributions in the DT–PWC are, in general, greater than two. Besides, we fully characterized the secrecy capacity in the low and high intensity regimes. In the low intensity regime, we derived a closed form expression for the secrecy capacity and observed that the it scales quadratically with the peak intensity constraint. Furthermore, we found that a binary input distribution with equiprobable mass points located at the origin and the peak intensity constraint is optimal. In the high intensity regime, we established that the secrecy capacity is upper bounded by a positive constant, thus, it does not scale with the peak intensity constraint. Finally, our numerical results indicated that under nonnegativity and peak intensity constraints, the secrecy capacity and the capacity of the DT–PWC channel cannot be obtained simultaneously in general, i.e., there is a tradeoff between the rate and its equivocation.
We start the proof by noting that for random variables $X$, $Y$ and $Z$ that form the Markov chain $X \rightarrow Y \rightarrow Z$, $I(X; Y| Z) = I(X; Y) - I(X; Z)$ and it is a concave function in $F_X$ [18, 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

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

where $p_{X_1}(x)$ and $p_{X_2}(x)$ are the probability density functions (PDF) of the random variables $X_1$ and $X_2$. Based on (48), we have the following Markov chain

$$Q \rightarrow X \rightarrow Y \rightarrow Z.$$  \hfill (49)

Following along the same lines as [18, Appendix A], one can show that

$$I(X; Y| Z, Q) - I(X; Y| Z) = -I(Q; Y| Z).$$

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 $Q \rightarrow X \rightarrow Y \rightarrow Z$, $I(X; Y| Z)$ is strictly concave in $F_X$, i.e., $I(Q; Y| Z) > 0$. Assume, to the contrary, that there exists an $F_X$ such that $I(Q; Y| Z) = 0$. This implies that random variables $Q$, $Y$ and $Z$ also form the Markov chain

$$Q \rightarrow Z \rightarrow Y.$$  \hfill (51)

Furthermore, from the Markov chain (49), we have

$$Q \rightarrow X \rightarrow Z.$$  \hfill (52)

Combining Markov chains (51) and (52) results in a new Markov chain given by

$$Q \rightarrow X \rightarrow Z \rightarrow Y.$$  \hfill (53)

Now, based on (49) and (53), we obtain the following

$$p(y, z, x) \bigg|_{\text{Markov chain (49)}} = p(y, z, x) \bigg|_{\text{Markov chain (53)}}$$

$$p_X(x) p(y|x) p(z|y) = p_X(x) p(z|x) p(y|z)$$

$$\frac{p(y|x)}{p(z|x)} = \frac{p(y|z)}{p(z|y)}.$$
We note that (54) holds for any \( y, z \in \mathbb{Z}_0^+ \) and \( x \in S_{F_X} \), where \( S_{F_X} \) is the support set of \( F_X \). As a result, for fixed values of \( y \) and \( z \) the right hand side (RHS) of (54) is fixed, while the left hand side (LHS) is a function of \( x \). Since \( Y|X \) and \( Z|X \) are Poisson distributed with mean \( x + \lambda_B \) and \( x + \lambda_E \), respectively, (54) reduces to

\[
e^{-\left(x + \lambda_B\right)} \frac{(x + \lambda_B)^y}{y!} \frac{(x + \lambda_E)^z}{z!} = \frac{p(y|z)}{p(z|y)}.
\]

(55)

To reach a contradiction, let us choose \( y = z = 1 \). Now, it is sufficient to show that the LHS of (55) is not a constant function in \( x \). To this end, let \( h(x) \) denote LHS (55) for \( y = z = 1 \). It is clear that \( h(x) \) is not a constant function in \( x \). For instance, \( h(0) = e^{(\lambda_E - \lambda_B) \frac{\lambda_B}{\lambda_E}} \neq e^{(\lambda_E - \lambda_B) \frac{\lambda_E}{2\lambda_E - \lambda_B}} = h(\lambda_E - \lambda_B) \) as \( \lambda_E > \lambda_B \). Therefore, we reach a contradiction. This, in turn, implies that \( I(Q;Y|Z) > 0 \) and as a result, \( I(X;Y|Z) \) is strictly concave in \( F_X \).

**APPENDIX B**

**SECRECY CAPACITY IN THE LOW INTENSITY REGIME**

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

\[
C_S = \frac{1}{2} \left[J_B(0) - J_E(0)\right] \max_{F_X \in \mathcal{A}^+} \text{Var}(X) + o(A^2).
\]

(56)

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 \( \text{Var}(X) \). Following along the same lines of [13, Appendix B], one can show that the input distribution \( F_X^* \in \mathcal{A}^+ \) that maximizes the \( \text{Var}(X) \) subject to nonnegativity and peak intensity constraint is a binary input distribution with mass points located at \( \{0,A\} \) with equal probabilities. Furthermore, \( \max_{F_X \in \mathcal{A}^+} \text{Var}(X) = \frac{A^2}{4} \).

**APPENDIX C**

**UPPER BOUND ON THE SECRECY CAPACITY IN THE HIGH INTENSITY REGIME**

We start the proof by noting that the DT–PWC is stochastically degraded due to the fact that \( \lambda_E > \lambda_B \). This implies that, the output of the eavesdropper’s channel \( Z \) can be written as \( Z = Y + N_D \), where \( N_D \) is a Poisson random variable with mean \( \lambda_D = \lambda_E - \lambda_B \). Therefore, \( H_Z(F_X^*) > H_{Z|N_D}(F_X^*) = H_Y(F_X^*) \) and consequently \( H_Z(F_X^*) > H_Y(F_X^*) \) for any nontrivial input distribution \( F_X^* \).
Furthermore, we can expand $H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X)$ as

$$H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X) = \mathbb{E}_{X,Z} \left[ -\log p_{Z|X}(z|x) \right] - \mathbb{E}_{X,Y} \left[ -\log p_{Y|X}(y|x) \right]$$

$$= (a) \mathbb{E}_{Z|X,Y} \left[ \mathbb{E}_{X,Y} \left[ \log p_{Y|X}(y|x) \right] \right] - \mathbb{E}_{Y|X,Z} \left[ \mathbb{E}_{X,Z} \left[ \log p_{Z|X}(z|x) \right] \right]$$

$$= \mathbb{E}_{X,Y,Z} \left[ \log \frac{p_{Y|X}(y|x)}{p_{Z|X}(z|x)} \right],$$

where \((a)\) follows as $\log p_{Y|X}(y|x)$ and $\log p_{Z|X}(z|x)$ do no depend on $Z$ and $Y$, respectively.

Plugging (2) and (3) into (57), we get

$$H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X) = \mathbb{E}_{X,Y,Z} \left[ \log \frac{e^{-(x+\lambda_B)}(x + \lambda_B)^y/y!}{e^{-(x+\lambda_E)}(x + \lambda_E)^z/z!} \right]$$

$$= \lambda_D + \mathbb{E}_X [(x + \lambda_B) \log (x + \lambda_B) - (x + \lambda_E) \log (x + \lambda_E)]$$

$$+ \mathbb{E}_{X,Y,Z} \left[ \log \frac{Z!}{Y!} \right],$$

Next, we consider the last term in (58) and try to find an upper bound on it. To this end, we first note that

$$\mathbb{E}_{X,Y,Z} \left[ \log \frac{Z!}{Y!} \right] = \mathbb{E}_X \left[ \mathbb{E}_{Y|X} \left[ \mathbb{E}_{Z|Y} \left[ \log \frac{Z!}{Y!} \right] \right] \right]$$

(59)
as $X \rightarrow Y \rightarrow Z$ is a Markov chain. Now, we have to find the conditional PDF of $Z|Z$. We proceed by observing that $Z = Y + N_D$, hence, one can show that

$$p_{Z|Y}(z|y) = \begin{cases} 0, & \text{if } z < y, \\ e^{-\lambda_D} \frac{(z-y)^{y}!}{(z-y)!}, & \text{if } z > y. \end{cases}$$

(60)

In what follows, we present chain of inequalities based on (60) which leads to the upper bound in (46):

$$\mathbb{E}_{X,Y,Z} \left[ \log \frac{Z!}{Y!} \right] = \mathbb{E}_X \left[ \mathbb{E}_{Y|X} \left[ \sum_{z=0}^{\infty} p_{Z|Y}(z|y) \log \frac{z!}{y!} \right] \right]$$

$$= \mathbb{E}_X \left[ \mathbb{E}_{Y|X} \left[ \sum_{z=0}^{\infty} e^{-\lambda_D} \frac{\lambda_D^{(z-y)}}{(z-y)!} \log \frac{z!}{y!} \right] \right]$$

$$= \mathbb{E}_X \left[ \mathbb{E}_{Y|X} \left[ \sum_{t=0}^{\infty} e^{-\lambda_D} \frac{\lambda_D^t}{t!} \log \frac{(t+y)!}{y!} \right] \right]$$

$$= \mathbb{E}_X \left[ \mathbb{E}_{Y|X} \left[ \sum_{t=0}^{\infty} e^{-\lambda_D} \frac{\lambda_D^t}{t!} \sum_{i=1}^{t} \log(y+i) \right] \right]$$
\[(b) \quad \mathbb{E}_X \left[ \sum_{t=0}^{+\infty} e^{-\lambda_D t} \frac{\lambda^t_D}{t!} \sum_{i=1}^{t} \log(x + \lambda_B + i) \right] \]
\[= \mathbb{E}_X \left[ \sum_{t=0}^{+\infty} e^{-\lambda_D t} \frac{\lambda^t_D}{t!} \left[ t \log(x + \lambda_B) + \sum_{i=1}^{t} \log \left( 1 + \frac{i}{x + \lambda_B} \right) \right] \right] \]
\[(c) \quad \leq \mathbb{E}_X \left[ \log(x + \lambda_B) \sum_{t=0}^{+\infty} e^{-\lambda_D t} \frac{\lambda^t_D}{t!} + \frac{1}{x + \lambda_B} \sum_{t=0}^{+\infty} e^{-\lambda_D t} \frac{\lambda^t_D}{t!} \frac{t(t+1)}{2} \right] \]
\[= \mathbb{E}_X \left[ \lambda_D \log(x + \lambda_B) + \frac{1}{x + \lambda_B} \left[ \frac{\lambda_D^2}{2} + \lambda_D \right] \right], \quad (61) \]

where \(b\) follows from sliding the expectation \(\mathbb{E}_{Y|X}\) through the summations and then applying the Jensen’s Inequality (as \(\log(y + i)\) is a concave function in \(y\)), and \(c\) follows from the fact that \(\log(1 + x) \leq x, \forall x \geq 0\). Now, using the upper bound in \((61)\), \(H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X)\) can be upper bounded as

\[
H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X) \leq \lambda_D + \mathbb{E}_X \left[ (x + \lambda_B) \log(x + \lambda_B) - (x + \lambda_E) \log(x + \lambda_E) \right] 
\]
\[+ \mathbb{E}_X \left[ \lambda_D \log(x + \lambda_B) + \frac{1}{x + \lambda_B} \left[ \frac{\lambda_D^2}{2} + \lambda_D \right] \right] \]
\[= \lambda_D + \mathbb{E}_X \left[ (x + \lambda_E) \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right] + \left[ \frac{\lambda_D^2}{2} + \lambda_D \right] \right] \mathbb{E}_X \left[ \frac{1}{x + \lambda_B} \right]. \quad (62) \]

Now, we note that since \(0 \leq x \leq A, \mathbb{E}_X \left[ \frac{1}{x + \lambda_B} \right] \leq \frac{1}{\lambda_B}\). Furthermore, denoting \(\psi(x) = (x + \lambda_E) \log \left[ \frac{x + \lambda_B}{x + \lambda_E} \right]\), we observe that \(\psi(x)\) is strictly negative, as \(\lambda_E > \lambda_B\), and it is an increasing function in \(x\) due to the fact that \(\frac{d\psi(x)}{dx} = -\log \left[ 1 + \frac{\lambda_D}{x + \lambda_B} \right] + \frac{\lambda_D}{x + \lambda_B} \geq 0\). This implies that the maximum value of \(\psi(x)\) is located at the end point of the interval \([0, A]\). Consequently, \(\psi(x) \leq (A + \lambda_E) \log \left[ \frac{A + \lambda_B}{A + \lambda_E} \right]\) and we have

\[
\psi(x) \leq \lim_{A \to \infty} (A + \lambda_E) \log \left[ \frac{A + \lambda_B}{A + \lambda_E} \right] = -\lambda_D. \quad (63) \]

From the upper bound on \(\mathbb{E}_X \left[ \frac{1}{x + \lambda_B} \right]\) and \((63)\), one can upper bound \((61)\) as

\[
H_{Z|X}(F^*_X) - H_{Y|X}(F^*_X) \leq \frac{\lambda_D^2}{2} + \lambda_D. \quad (64) \]

This completes the proof of the proposition.
REFERENCES


