# The Continuous-Time Poisson Channel Has Infinite Covert Communication Capacity

Ligong Wang
ETIS—Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS
Cergy-Pontoise, France
ligong.wang@ensea.fr

***Abstract*—We consider the problem of communication over a continuous-time, infinite-bandwidth Poisson channel without peak-power constraint, but subject to a covertness constraint: the relative entropy between the output distributions when a codeword is transmitted and when no input is provided to the channel must tend to zero as total communication time grows large. We show that, under such a constraint, the capacity of this channel, in nats per second, is infinity.**

## I. Introduction

Covert communication, or communication with low probability of detection [1]–[4], refers to applications where the transmitter and the receiver must keep an eavesdropper from discovering the fact that they are using the channel to communicate. Specifically, the signals observed by the eavesdropper when the channel is used must be statistically close to the signals observed when the channel is not used. For additive white Gaussian noise (AWGN) channels, the channel not being used is modeled by the transmitter always sending zero; for a discrete memoryless channel (DMC), this is modeled by the transmitter sending an "innocent" symbol—a specific input symbol whose value is given as part of the channel model. For a DMC, if the output distribution at the eavesdropper generated by the innocent symbol is a convex combination of the output distributions generated by the other, non-innocent input symbols, then a positive covert communication rate is achievable; otherwise the maximum amount of information that can be covertly communicated grows like the square root of the total number of channel uses [3]. For the AWGN channel, the latter, square-root scaling law applies [1], [3].

As most practically relevant channel models appear to follow the square-root scaling law and thus have zero capacity (in nats per channel use)[1] for covert communication, some works have considered variations of the DMC and AWGN channel models. Among them, [5]–[7] consider channels with parameters that are unknown to the eavesdropper, and [8] considers channels with channel-state information (CSI) available to the transmitter. These works show that, under such additional assumptions, positive covert communication rates are sometimes achievable.

The current paper is concerned with covert communication over a continuous-time channel: the Poisson channel [9]–[11]. The input to such a channel is a continuous-time, nonnegative waveform, and the output is a doubly stochastic Poisson process whose conditional intensity is given by the input waveform plus a constant describing background noise. Unlike [9]–[11], we impose *no peak constraint* on the input waveform. We assume that, when the transmitter is not sending a message, it sends the all-zero waveform. For simplicity, we assume that the intended receiver and the eavesdropper observe the same channel output, but it will become apparent from our proof that our results also hold when they observe different signal attenuations and noise levels, as long as the length of the secret key to be used is not a concern. We shall show that the Poisson channel has an infinite per-second covert communication capacity. Thus, the continuous-time Poisson channel follows no square-root law that is analogous to the one for discrete-time channels; in contrast, it allows the number of covert information nats to grow super-linearly with total communication time.

Compared to previously studied channel models that allow positive-rate covert communication, our model is of a different nature. First, the Poisson channel is memoryless, but is dissimilar to those DMCs that allow positive-rate covert communication [3], in the sense that its output distribution conditional on the all-zero input is *not* a convex combination of the output distributions conditional on any input waveforms that are not almost everywhere zero. To see this, simply note that any input waveform with a positive Lebesgue integral will increase the expectation of the total number of output arrivals compared to the all-zero input waveform. Second, our model is not related to the queueing channel considered in [12]. In [12] it is assumed that, when no message is being communicated via the queueing channel, packets arrive in a random, Poisson process, making the problem similar to that of "stealth" [13]. Third, we do not assume any unknown parameter or channel state as in [5]–[8]. Such additional assumptions would allow the transmitter to send a non-vanishing average signal power or a non-vanishing percentage of non-innocent symbols. As we shall see, our Poisson model requires that the average input power tend to zero as total communication time grows large, but even this vanishing input power provides unbounded communication capacity.

In some sense, our model allows infinite covert communication capacity because it is a continuous-time channel of infinite bandwidth. Indeed, once we restrict the communication

[1]In some works, the name "covert capacity" refers to the scaling constant of information nats with respect to the square root of total number of channel uses. In the current work, we use "capacity" is its traditional sense.

bandwidth to a certain value to reduce the Poisson channel to discrete time, we can use existing capacity results such as those in [14], [15] to show that the maximum covert communication rate vanishes as total communication time grows large. Also crucial is the absence of a peak-power constraint. It is well known that a peak-constrained infinite-bandwidth Poisson channel has a finite per-second communication capacity [9]–[11], whereas without a peak constraint, its capacity for standard, non-covert communication is infinity. One can extend these classic results to show that, when subject to a peak constraint, the Poisson channel again has zero covert communication capacity. We shall not further discuss these cases in the current paper.

Our signaling scheme to achieve infinite covert communication rates is to divide the total communication time into short slots, and use independent on-off signaling within these slots. We let the length of the slots decrease, and the value of the "on" signal increase, as total communication time grows large.

The rest of this paper is arranged as follows. Section II formulates the problem and states the result. Section III contains the proof of the result, but with some steps omitted or shortened. Section IV then concludes the paper with some discussions.

## II. Problem Formulation and Result

Consider a continuous-time Poisson channel with constant dark current $\lambda > 0$. The input to the channel must be a Lebesgue-measurable, nonnegative signal $x(t) \geq 0$, $t \in [0, T]$. Conditional on this input signal, the channel output $Y(t)$, $t \in [0, T]$, is a Poisson process whose time-$t$ intensity is $x(t) + \lambda$.

A codebook on $[0, T]$ at rate $R$ nats per second for communication over this channel contains $e^{TR}$ input signals $x_m(t)$, $t \in [0, T]$, $m \in \{1, \ldots, e^{TR}\}$. An encoder maps a message $m \in \{1, \ldots, e^{TR}\}$ to the corresponding input signal $x_m(t)$, $t \in [0, T]$, and sends this signal into the channel. A corresponding decoder maps the output signal $y(t)$, $t \in [0, T]$, to the decoded message $\hat{m} \in \{1, \ldots, e^{TR}\}$. We allow the transmitter and the receiver to use a *random code* to communicate their message. The eavesdropper is assumed to know the distribution according to which the random code is chosen, but not the actual choice.

We do *not* impose a peak constraint on the input signal, hence the value of $x_m(t)$ at a specific time $t$ may be arbitrarily large. As we shall see, our result does not depend on whether there is an average-power constraint or not, i.e., it holds under any upper bound on the value of

$$\frac{1}{T} \mathsf{E}\left[\int_0^T x_M(t)\,\mathrm{d}t\right],$$

where $M$ is uniformly chosen from $\{1, \ldots, e^{TR}\}$.

Let $Q_T$ denote the distribution of $Y(t)$, $t \in [0, T]$, when there is no input signal, i.e., when $x(t) = 0$ for all $t \in [0, T]$. Thus, $Q_T$ describes the homogeneous Poisson process of intensity $\lambda$ on $[0, T]$. Further, let $P_T$ denote the distribution of $Y(t)$, $t \in [0, T]$, induced by the random codebook and the randomly chosen message $M$. Thus, $P_T$ is the unconditional output distribution when the channel is used to communicate a message.

We say that a rate $R$ is *achievable covertly* if, for every $T > 0$, we can construct a random code on $[0, T]$ of rate at least $R$ such that

1) the probability of a decoding error, i.e., of $\hat{M} \neq M$, tends to zero as $T$ tends to infinity; and
2) $D(P_T \| Q_T)$ tends to zero as $T$ tends to infinity.

Note that, by Pinsker's inequality [16], the second requirement implies that the total variation distance between $P_T$ and $Q_T$ must also tend to zero as $T \to \infty$, hence there will be no non-trivial test to distinguish $P_T$ and $Q_T$. We then define the *covert communication capacity* for this channel as the supremum of all rates that are achievable covertly.

Our main result is the following theorem.

*Theorem 1:* The covert communication capacity of the continuous-time Poisson channel described above is infinity, i.e., all positive rates are achievable covertly.

## III. Proof of Theorem 1

Clearly, only the achievability part of the theorem needs proof. To this end, we describe a signaling scheme and analyze its performance.

We divide the interval $[0, T]$ into slots of $\tau$ seconds.[2] We choose $X(t)$ to be constant within each slot, and independent and identically distributed (IID) between different slots; in particular, within every slot, $X(t) = X'$ for all $t$ in the slot, where

$$X' = \begin{cases} A, & \text{with probability } q, \\ 0, & \text{with probability } 1 - q. \end{cases} \tag{1}$$

Here, $\tau$, $A$, and $q$ all depend on $T$. We choose them as follows:

$$\tau = T^{-1} e^{-T} \tag{2a}$$
$$A = e^{T/2} \tag{2b}$$
$$q = T^{-3/4} e^{-T/2}. \tag{2c}$$

We note that the average input power by the above choice equals

$$qA = T^{-3/4}, \tag{3}$$

which tends to zero as $T$ tends to infinity. Hence any average-power constraint on the input will be satisfied when $T$ is large enough.

Our analysis of this scheme is divided into two parts. The first part shows that the scheme is covert. The second part computes the rates that can be achieved by this scheme.

[2]We ignore the issue that $T$ may not be divisible by $\tau$, because it only has an edge effect, which vanishes when $T \to \infty$.

### A. Covertness analysis

Within each $\tau$-second slot, let $Y'$ denote the total number of output arrivals within this slot. Note that, since with probability one $X(t)$ is constant within the slot (both when the transmitter is sending and when it is not sending a message), $Y'$ forms a sufficient statistic of the output process in the slot, both for the receiver to decode the message and for the eavesdropper to detect communication activity. When no message is being communicated (i.e., under $Q_T$), $Y'$ has the Poisson distribution of mean $\lambda\tau$; when the transmitter is sending a message (under $P_T$), with probability $1-q$ the random variable $Y'$ has the Poisson distribution of mean $\lambda\tau$, and with probability $q$ it has the Poisson distribution of mean $(A+\lambda)\tau$. Let these two distributions of $Y'$ be denoted $Q'$ and $P'$, respectively. Since different slots are independent under both $P_T$ and $Q_T$, we have

$$D(P_T\|Q_T) = \frac{T}{\tau} D(P'\|Q'). \tag{4}$$

We next compute $D(P'\|Q')$. For each $i \in \mathbb{Z}_0^+$, we have

$$Q'(i) = \frac{e^{-\lambda\tau}\lambda^i}{i!} \tag{5a}$$

$$P'(i) = (1-q)\frac{e^{-\lambda\tau}\lambda^i}{i!} + q\frac{e^{-(A+\lambda)\tau}(A+\lambda)^i}{i!}. \tag{5b}$$

Hence

$$D(P'\|Q') = \sum_{i=0}^{\infty} P'(i)\log\frac{P'(i)}{Q'(i)} \tag{6}$$

$$= \sum_{i=0}^{\infty} P'(i)\log\left(1-q+qe^{-A\tau}\left(1+\frac{A}{\lambda}\right)^i\right). \tag{7}$$

Denote the summand on the right-hand side of (7) for $i$ by $\sigma_i$. We bound them separately.

For $\sigma_0$ we have

$$P'(0) = (1-q)e^{-\lambda\tau} + qe^{-(A+\lambda)\tau} \tag{8}$$

$$\geq (1-q)(1-\lambda\tau) + q(1-(A+\lambda)\tau) \tag{9}$$

$$= 1-(\lambda+qA)\tau, \tag{10}$$

where we used $e^{-a} \geq 1-a$ for all $a \in \mathbb{R}$. We also have

$$\log(1-q+qe^{-A\tau}) \leq -q(1-e^{-A\tau}) \tag{11}$$

$$\leq -q\left(A\tau - \frac{A^2\tau^2}{2}\right), \tag{12}$$

where the first inequality follows because $\log(1+a) \leq a$ for all $a \in \mathbb{R}$, and the second inequality because $e^{-a} \leq 1-a+\frac{a^2}{2}$ for $a \geq 0$. Note that the left-hand side of (11) is non-positive, whereas, by our choices (2), the right-hand side of (10) is positive for large enough $T$. We can hence combine (10) and (12) to obtain, for large enough $T$,

$$\sigma_0 \leq -(1-(\lambda+qA)\tau)\cdot q\left(A\tau - \frac{A^2\tau^2}{2}\right) \tag{13}$$

$$\leq -qA\tau\left(1-\left(\lambda+qA+\frac{A}{2}\right)\tau\right), \tag{14}$$

where the last inequality follows by dropping a negative term.

For $\sigma_1$ we have

$$P'(1) = (1-q)\lambda\tau\underbrace{e^{-\lambda\tau}}_{\leq 1} + q(A+\lambda)\tau\underbrace{e^{-(A+\lambda)\tau}}_{\leq 1} \tag{15}$$

$$\leq (\lambda+qA)\tau \tag{16}$$

and

$$\log\left(1-q+qe^{-A\tau}\left(1+\frac{A}{\lambda}\right)\right)$$

$$\leq 1-q+qe^{-A\tau}\left(1+\frac{A}{\lambda}\right)-1 \tag{17}$$

$$= q\left(\underbrace{e^{-A\tau}}_{\leq 1}\left(1+\frac{A}{\lambda}\right)-1\right) \tag{18}$$

$$\leq \frac{qA}{\lambda}. \tag{19}$$

Thus

$$\sigma_1 \leq (\lambda+qA)\tau\cdot\frac{qA}{\lambda} = qA\tau\left(1+\frac{qA}{\lambda}\right). \tag{20}$$

For the remaining summands where $i \geq 2$, we have

$$\log\left(1-q+q\underbrace{e^{-A\tau}}_{\leq 1}\left(1+\frac{A}{\lambda}\right)^i\right)$$

$$\leq \log\left(1+\underbrace{q}_{\leq 1}\left(\left(1+\frac{A}{\lambda}\right)^i-1\right)\right) \tag{21}$$

$$\leq \log\left(1+\left(1+\frac{A}{\lambda}\right)^i-1\right) \tag{22}$$

$$= i\log\left(1+\frac{A}{\lambda}\right) \tag{23}$$

$$\leq i\cdot\frac{A}{\lambda}. \tag{24}$$

Thus

$$\sum_{i=2}^{\infty}\sigma_i \leq \frac{A}{\lambda}\sum_{i=2}^{\infty}P'(i)\cdot i, \tag{25}$$

where

$$\sum_{i=2}^{\infty}P'(i)\cdot i$$

$$= \sum_{i=1}^{\infty}P'(i)\cdot i - P'(1)\cdot 1 \tag{26}$$

$$= \mathsf{E}_{P'}[Y'] - P'(1) \tag{27}$$

$$= \{(1-q)\lambda\tau + q(A+\lambda)\tau\}$$

$$- \{(1-q)\underbrace{e^{-\lambda\tau}}_{\geq 1-\lambda\tau}\lambda\tau + q\underbrace{e^{-(A+\lambda)\tau}}_{\geq 1-(A+\lambda)\tau}(A+\lambda)\tau\} \tag{28}$$

$$\leq (1-q)\lambda\tau + q(A+\lambda)\tau - (1-q)(1-\lambda\tau)\lambda\tau$$

$$- q(1-(A+\lambda)\tau)(A+\lambda)\tau \tag{29}$$

$$= (1-q)\lambda^2\tau^2 + q(A+\lambda)^2\tau^2. \tag{30}$$

Hence we obtain

$$\sum_{i=2}^{\infty} \sigma_i \le \frac{A}{\lambda}\big((1-q)\lambda^2\tau^2 + q(A+\lambda)^2\tau^2\big). \tag{31}$$

Combining (4), (7), (14), (20), and (31), we have, for large enough $T$,

$$\begin{aligned} &D(P_T\|Q_T) \\ &\le \frac{T}{\tau}\bigg\{ -qA\tau\bigg(1-\bigg(\lambda+qA+\frac{A}{2}\bigg)\tau\bigg) + qA\tau\bigg(1+\frac{qA}{\lambda}\bigg) \\ &\qquad + (1-q)\lambda A\tau^2 + \frac{qA}{\lambda}(A+\lambda)^2\tau^2\bigg\} \qquad (32) \\ &= \bigg\{\lambda A\tau + q^2A^2\tau + \frac{qA^2\tau}{2} + \frac{q^2A^2}{\lambda} + \frac{qA}{\lambda}(A+\lambda)^2\tau\bigg\}\cdot T. \qquad (33) \end{aligned}$$

Replacing $\tau$, $A$, and $q$ by their values in (2) in the above implies that

$$\lim_{T\to\infty} D(P_T\|Q_T) = 0. \tag{34}$$

Hence the random code we generate is covert in the limit where $T \to \infty$.

### *B. Communication Rate*

To bound the communication rates achievable using the proposed input distribution, we consider suboptimal decoding schemes in which each output $Y'$ is mapped to a binary random variable $\hat{Y}$:

$$\hat{Y} = \begin{cases} 0, & Y' = 0 \\ 1, & \text{otherwise.} \end{cases} \tag{35}$$

Then we have

$$\Pr(\hat{Y}=1|X'=0) = 1-e^{-\lambda\tau} \tag{36a}$$
$$\Pr(\hat{Y}=1|X'=A) = 1-e^{-(A+\lambda)\tau}. \tag{36b}$$

We then use the information spectrum method [17], [18] to lower-bound the achievable rates. Let the input vector of length $T/\tau$ be denoted $\mathbf{X}'$ and the corresponding (binary) output vector be denoted $\hat{\mathbf{Y}}$, then the following rate is achievable:

$$P\text{-}\liminf_{T\to\infty} \frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}, \tag{37}$$

where $P$-$\liminf$ denotes the limit infimum in probability; see [17], [18]. To evaluate (37), we compute the mean and the variance of the random variable in the $P$-$\liminf$.

For any fixed $T$, $(\mathbf{X}', \hat{\mathbf{Y}})$ is IID across different time slots, hence

$$\mathsf{E}\bigg[\frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}\bigg] = \frac{1}{\tau}I(X';\hat{Y}). \tag{38}$$

By (1) and (36), $I(X';\hat{Y})$ can be written as

$$\begin{aligned} I(X';\hat{Y}) &= H_{\mathrm{b}}\Big((1-q)(1-e^{-\lambda\tau}) + q(1-e^{-(\lambda+A)\tau})\Big) \\ &\quad - (1-q)H_{\mathrm{b}}(1-e^{-\lambda\tau}) - qH_{\mathrm{b}}(1-e^{-(\lambda+A)\tau}), \end{aligned} \tag{39}$$

where $H_{\mathrm{b}}(\cdot)$ denotes the binary entropy:

$$H_{\mathrm{b}}(p) = -p\log p - (1-p)\log(1-p), \quad p\in[0,1]. \tag{40}$$

We bound the three terms on the right-hand side of (39) separately.

For the first term on the right-hand side of (39), we drop a nonnegative term to get

$$\begin{aligned} &H_{\mathrm{b}}\Big((1-q)(1-e^{-\lambda\tau}) + q(1-e^{-(\lambda+A)\tau})\Big) \\ &\quad \ge \Big((1-q)(1-e^{-\lambda\tau}) + q(1-e^{-(\lambda+A)\tau})\Big) \\ &\qquad \cdot \log\frac{1}{(1-q)(1-e^{-\lambda\tau}) + q(1-e^{-(\lambda+A)\tau})}. \end{aligned} \tag{41}$$

Then we bound

$$\begin{aligned} &(1-q)\underbrace{(1-e^{-\lambda\tau})}_{\ge\lambda\tau-\frac{\lambda^2\tau^2}{2}} + q\underbrace{(1-e^{-(\lambda+A)\tau})}_{\ge(\lambda+A)\tau-\frac{(\lambda+A)^2\tau^2}{2}} \\ &\ge (1-q)\bigg(\lambda\tau - \frac{\lambda^2\tau^2}{2}\bigg) + q\bigg((\lambda+A)\tau - \frac{(\lambda+A)^2\tau^2}{2}\bigg) \qquad (42) \\ &\ge \bigg((\lambda+qA) - \frac{q(\lambda+A)^2\tau}{2} - \frac{\lambda^2\tau}{2}\bigg)\tau, \qquad (43) \end{aligned}$$

where the last inequality follows by dropping a term $\frac{q\lambda^2\tau^2}{2}$. We also have

$$\begin{aligned} &(1-q)\underbrace{(1-e^{-\lambda\tau})}_{\le\lambda\tau} + q\underbrace{(1-e^{-(\lambda+A)\tau})}_{\le(\lambda+A)\tau} \\ &\le (1-q)\lambda\tau + q(\lambda+A)\tau = (\lambda+qA)\tau. \end{aligned} \tag{44}$$

Combining (41), (43), and (44) we have, for large enough $T$,

$$\begin{aligned} &H_{\mathrm{b}}\Big((1-q)(1-e^{-\lambda\tau}) + q(1-e^{-(\lambda+A)\tau})\Big) \\ &\ge \bigg((\lambda+qA) - \frac{q(\lambda+A)^2\tau}{2} - \frac{\lambda^2\tau}{2}\bigg)\tau\log\frac{1}{(\lambda+qA)\tau}. \end{aligned} \tag{45}$$

By our choices (2), for large enough $T$, $\lambda\tau < 1/2$ and $(\lambda+A)\tau < 1/2$. Since $H_{\mathrm{b}}(\cdot)$ is monotonically increasing on $(0, \frac{1}{2})$, we can bound

$$\begin{aligned} H_{\mathrm{b}}(1-e^{-\lambda\tau}) &\le H_{\mathrm{b}}(\lambda\tau) \qquad (46) \\ &= \lambda\tau\log\frac{1}{\lambda\tau} + (1-\lambda\tau)\underbrace{\log\frac{1}{1-\lambda\tau}}_{\le\frac{\lambda\tau}{1-\lambda\tau}} \qquad (47) \\ &\le \lambda\tau\log\frac{1}{\lambda\tau} + \lambda\tau, \qquad (48) \end{aligned}$$

and similarly,

$$H_{\mathrm{b}}(1-e^{-(A+\lambda)\tau}) \le (A+\lambda)\tau\log\frac{1}{(A+\lambda)\tau} + (A+\lambda)\tau. \tag{49}$$

Combining (38), (39), (45), (48), and (49) yields, for large enough $T$,

$$\begin{aligned}
&\mathsf{E}\left[\frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}\right] \\
&\geq \left((\lambda+qA)-\frac{q(\lambda+A)^2\tau}{2}-\frac{\lambda^2\tau}{2}\right)\log\frac{1}{(\lambda+qA)\tau} \\
&\quad -(1-q)\lambda\log\frac{1}{\lambda\tau}-(1-q)\lambda \\
&\quad -q(A+\lambda)\log\frac{1}{(A+\lambda)\tau}-q(A+\lambda) \qquad (50)\\
&=(1-q)\lambda\log\frac{\lambda}{\lambda+qA}+q(\lambda+A)\log\frac{\lambda+A}{\lambda+qA} \\
&\quad -(\lambda+qA)-\frac{q(\lambda+A)^2+\lambda^2}{2}\tau\log\frac{1}{(\lambda+qA)\tau}. \qquad (51)
\end{aligned}$$

One can verify that, for our choices (2), for large enough $T$, the right-hand side of (51) is dominated by its second term, which behaves like

$$\begin{aligned}
&\mathsf{E}\left[\frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}\right] \\
&\approx q(\lambda+A)\log\frac{\lambda+A}{\lambda+qA} \qquad (52)\\
&=T^{-3/4}e^{-T/2}(\lambda+e^{T/2})\log\frac{\lambda+e^{T/2}}{\lambda+T^{-3/4}e^{-T/2}e^{T/2}} \qquad (53)\\
&\approx T^{-3/4}\left(\frac{T}{2}-\log\lambda\right)\approx\frac{T^{1/4}}{2}, \qquad (54)
\end{aligned}$$

which tends to infinity as $T\to\infty$.

We briefly outline our bound on the variance of the random variable inside the $P$-$\liminf$ in (37). Since $(\mathbf{X}',\hat{\mathbf{Y}})$ is IID across different slots, we have

$$\begin{aligned}
&\mathsf{Var}\left[\frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}\right] \\
&\qquad =\frac{1}{T^2}\cdot\frac{T}{\tau}\,\mathsf{Var}\left[\log\frac{P_{\hat{Y}|X'}(\hat{Y}|X')}{P_{\hat{Y}}(\hat{Y})}\right] \qquad (55)\\
&\qquad \leq\frac{1}{T\tau}\,\mathsf{E}\left[\left(\log\frac{P_{\hat{Y}|X'}(\hat{Y}|X')}{P_{\hat{Y}}(\hat{Y})}\right)^2\right]. \qquad (56)
\end{aligned}$$

For large enough $T$ and for our choices (2), one can check that the second moment on the right-hand side of (56) is dominated by one term:

$$\mathsf{E}\left[\left(\log\frac{P_{\hat{Y}|X'}(\hat{Y}|X')}{P_{\hat{Y}}(\hat{Y})}\right)^2\right]\approx qA\tau(\log A)^2=\frac{T^{5/4}}{4}\tau, \qquad (57)$$

thus

$$\mathsf{Var}\left[\frac{1}{T}\log\frac{P_{\hat{\mathbf{Y}}|\mathbf{X}'}(\hat{\mathbf{Y}}|\mathbf{X}')}{P_{\hat{\mathbf{Y}}}(\hat{\mathbf{Y}})}\right]\lesssim\frac{T^{1/4}}{4}. \qquad (58)$$

Combining (54) and (58) and applying Chebyshev's inequality yield that (37) is infinity, and complete the proof.

## IV. Discussions

We have shown that the continuous-time Poisson channel with neither peak nor bandwidth constraint allows unbounded data rates for covert communication. To demonstrate this, we choose both peak power and bandwidth to grow exponentially with total communication time.

Our work shows that the square-root law for covert communication over discrete-time channels does not apply to continuous-time channels without a bandwidth constraint. This can also be seen via a recent work [19] in which, using different techniques, we show that a Gaussian channel without bandwidth constraint also permits positive-rate covert communication. An interesting direction for future work may be to construct a general framework for covert communication over continuous-time channels.

## References


[1] B. A. Bash, D. Goekel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Select. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sept. 2013.

[2] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. IEEE Int. Symp. Inform. Theory*, Istanbul, Turkey, July 10–15 2013.

[3] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Trans. Inform. Theory*, vol. 62, no. 6, pp. 3493–3503, June 2016.

[4] M. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inform. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[5] P. H. Che, M. Bakshi, C. Chan, and S. Jaggi, "Reliable deniable communication with channel uncertainty," in *Proc. Inform. Theory Workshop (ITW)*, Hobart, Australia, Nov. 2–5, 2014.

[6] S. Lee, R. Baxley, M. Weitnauer, and B. Walkenhorst, "Achieving undetectable communication," *IEEE J. Select. Topics. Signal Processing.*, vol. 9, no. 7, pp. 1195–1205, Oct. 2015.

[7] T. Sobers, B. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Trans. Wireless Comm.*, vol. 16, no. 9, pp. 6193–6206, 2017.

[8] S.-H. Lee, L. Wang, A. Khisti, and G. W. Wornell, "Covert communication with noncausal channel-state information at the transmitter," in *Proc. IEEE Int. Symp. Inform. Theory*, Aachen, Germany, June 25–30 2017.

[9] Y. Kabanov, "The capacity of a channel of the Poisson type," *Theory of Probability and Its Appl.*, vol. 23, pp. 143–147, 1978.

[10] M. H. A. Davis, "Capacity and cutoff rate for Poisson-type channels," *IEEE Trans. Inform. Theory*, vol. 26, pp. 710–715, Nov. 1980.

[11] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel — parts I and II," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1462–1471, Nov. 1988.

[12] P. Mukherjee and S. Ulukus, "Covert bits through queues," in *IEEE Conf. Comm. and Network Security*, Philadelphia, PA, USA, Oct. 2016.

[13] J. Hou and G. Kramer, "Effective secrecy: reliability, confusion and stealth," in *Proc. IEEE Int. Symp. Inform. Theory*, Honolulu, HI, USA, June 29–July 4 2014.

[14] A. Lapidoth, J. H. Shapiro, V. Venkatesan, and L. Wang, "The discrete-time Poisson channel at low input powers," *IEEE Trans. Inform. Theory*, vol. 57, no. 6, pp. 3260–3272, June 2011.

[15] L. Wang and G. W. Wornell, "A refined analysis of the Poisson channel in the high-photon-efficiency regime," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 4299–4311, July 2014.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: John Wiley & Sons, 2006.

[17] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1147–1157, July 1994.

[18] T. S. Han, *Information Spectrum Methods in Information Theory*. Springer Verlag, 2003.

[19] L. Wang, "On covert communication over infinite-bandwidth Gaussian channels," 2018, to be presented at *SPAWC*.