Finite Blocklength Analysis of the Continuous-Time Poisson Channel

Dina Abdelhadi

Advisor: Stefan Moser
Co-Advisor: Amos Lapidoth, Ligong Wang
Acknowledgments

I would like to express my gratitude to Professor Stefan Moser for the guidance, advice and encouragement that he provided throughout this project. I am also grateful to Professor Amos Lapidoth and Dr. Ligong Wang for the time they devoted for discussions.

Zürich, 12 June 2017

Dina Abdelhadi
Abstract

Information theory’s most celebrated result is Shannon’s formula for capacity $C = \max_{P_X} I(X;Y)$. It gives an upper bound on the rate that can be transmitted reliably over a channel. There is just one caveat; ‘reliably’ means that the probability of error tends to zero as the blocklength $n$ approaches infinity. But what if we were to ask the more practical question: How much information can be transmitted over a channel if the blocklength is not infinite? In this case, we set some probability of error that should not be exceeded. This question has been the focus of many works which mainly addressed discrete-time channels. In this report, we extend these results slightly into the arena of continuous-time channels, where, instead of a finite blocklength, we deal with a waveform of finite duration. We study the continuous-time Poisson channel under the assumptions of a pulse amplitude modulated waveform in the limit of the pulse-width approaching 0.
Contents

1 Introduction
   1.1 The Poisson Channel ........................................ 1
   1.2 Problem Statement .......................................... 1
   1.3 Discretized Model of the Poisson Channel ................. 2
   1.4 Key Ideas and Report Organization .......................... 3
      1.4.1 Deriving Bounds on $M$ from Bounds on $\epsilon$ .... 3
      1.4.2 Report Organization .................................... 4

2 Definitions
   2.1 Notation ....................................................... 5
   2.2 Channel Quantities .......................................... 5
   2.3 Other Definitions ........................................... 6

3 Upper Bound
   3.1 Lower Bounding Poisson Channel Error Probability ....... 7
   3.2 Upper Bound on $\log M$ ..................................... 9
   3.3 Region of Validity .......................................... 10

4 Achievability Bounds
   4.1 Achievability using the RCU Bound ......................... 11
      4.1.1 Proof Ingredients ...................................... 12
      4.1.2 Achievability Bound Proof ............................ 15
      4.1.3 Achievability in Continuous Time as $\Delta \to 0$ ..... 17
   4.2 Achievability Bound Using Dependence Testing Bound .... 17
      4.2.1 Region of Validity .................................... 18

5 Result and Conclusion ........................................... 19
A Appendix

A.1 Conditional Variance and Third Order Absolute Moment of Channel (1.1) 21

Bibliography 23

List of Figures 25
Chapter 1

Introduction

1.1 The Poisson Channel

The Poisson channel is used to model optical communication channels with direct detection, where the input to the channel is a continuous signal intensity \( \lambda(t) \geq 0 \). The output of the channel is a Poisson counting process modeling the number of photons \( \nu(t) \) arriving in a time interval \( \tau \) as follows

\[
P[\nu(t + \tau) - \nu(t) = j] = \frac{\Lambda^j e^{-\Lambda}}{j!}, \quad j \in \mathbb{N}_0,
\]

where

\[
\Lambda = \int_t^{t+\tau} (\lambda(t') + \lambda_0) \, dt',
\]

and \( \lambda_0 \) is a nonnegative constant called the dark current, expressing noise present due to the random generation of charges.

1.2 Problem Statement

We wish to analyze the finite blocklength behaviour of the continuous-time Poisson channel, with

- a peak power constraint \( \max \lambda(t) = A \) and
- an average power constraint

\[
\frac{1}{T} \int_0^T \lambda(t) \, dt \leq \sigma A,
\]

where \( 0 \leq \sigma \leq 1 \).
More precisely, we would like to obtain upper and lower bounds on the number of transmissible codewords $M$ over the Poisson channel during a finite duration $T$ while maintaining an average error probability not larger than $\epsilon_{\text{avg}}$.

### 1.3 Discretized Model of the Poisson Channel

In the continuous-time case, a finite blocklength corresponds to a finite duration $T$. We assume a pulse amplitude modulated waveform input of a pulse-width $\Delta$. During a time window $T$, the channel is used $n = T/\Delta$ times. In order to obtain a discretized model of the continuous-time Poisson channel, the assumptions made in [1] and justified in [2] are adopted. These assumptions have been shown in [2] to only degrade performance in a negligible way as $\Delta \to 0$ for any finite $T$. Thus, making the following assumptions as $\Delta \to 0$ implies no increase in the probability of error:

- samples at discrete time instants are sufficient statistics
  \[ y_i = \nu((i + 1)\Delta) - \nu(i\Delta), \]

- the probability that two photons arrive during the same pulse duration is negligible when $\Delta \to 0$. Thus, the output distribution is binary, where the detector interprets the event of arrival of more than one photon to be the same as the arrival of none, i.e., $y_i \in \{0, 1\}$,

- and the capacity achieving input distribution is binary, i.e., $\lambda(t) \in \{0, A\}$.

Thus, in the limit of $\Delta \to 0$, the Poisson channel is equivalent to a discrete memoryless binary-input, binary-output channel

\[
W_p(y|x) = \begin{bmatrix}
    w(0|0) & w(0|1) \\
    w(1|0) & w(1|1)
\end{bmatrix} = \begin{bmatrix}
    1 - A \Delta e^{-\lambda_0 s} & 1 - A \Delta e^{-\lambda_0 A (s+1)} (s+1) \\
    A \Delta e^{-\lambda_0 A s} & A \Delta e^{-\lambda_0 A (s+1)} (s+1)
\end{bmatrix},
\]

where $s = \lambda_0/A$. As defined in [1], whenever $\lambda(t) = A$ during some interval $((n-1)\Delta, n\Delta]$, $x_n$ is taken to be $1$, otherwise $\lambda(t) = x_n = 0$. 

2
The capacity achieving input distribution is given in [1] by
\[ q = P[X = 1] = \min \left\{ \frac{(1+s)^{1+s}}{s^re} - s, \sigma \right\}, \]
where this input distribution takes care of the average power constraint. Thus, if we obtain achievability bounds for an arbitrary input distribution, we can simply plug in \( q \), and this way the average power constraint is sure to have been satisfied.

1.4 Key Ideas and Report Organization

In this report, we mainly extend the finite blocklength results of [3] which were derived for discrete memoryless channels into the realm of continuous-time channels. We use the simplified model of a binary-input binary-output Poisson model. The aspects arising due to the continuous-time nature of the channel exhibit themselves through the pulse-width \( \Delta \), upon which the channel transition probability matrix depends. We apply the same bounds for the DMC, but as a function of \( \Delta \), then take the limit as \( \Delta \to 0 \).

1.4.1 Deriving Bounds on \( M \) from Bounds on \( \epsilon \)

A key technique used in [3] is to derive bounds on \( \log M \) using bounds on error probability. For example, consider the case when we have an expression for an upper bound on error probability as a function of \( M \),
\[ \epsilon(M) \leq f(M). \]
We wish to extract from this bound an achievability bound for \( M \) as a function of \( \epsilon \). The idea is to set some constant \( \epsilon_1 \) and find \( M^* \) small enough such that the error as a function of \( M^* \) satisfies \( \epsilon(M^*) \leq \epsilon_1 \). More formally, we find \( M^* \) such that
\[ \epsilon(M^*) \leq \epsilon_1, \]
and solve for \( M^* \) to find an expression for an achievable \( M^* \)
\[ f(M^*) = \epsilon_1 - \delta \]
\[ M^* = f^{-1}(\epsilon_1 - \delta). \]
Chapter 1

Note that $M^*$ does indeed satisfy the requirement on error probability

$$\epsilon(M^*) \leq f(M^*) = f(f^{-1}(\epsilon_1 - \delta)) = \epsilon_1 - \delta.$$ 

The same strategy can be used to derive upper bounds on $\log M$.

1.4.2 Report Organization

The rest of this report is organized as follows: Chapter 2 includes some definitions and hints on notation that will be useful throughout the report and in Chapter 3 an upper bound on $\log M(T, \epsilon)$ of the Poisson channel is obtained. In Chapter 4, achievability bounds of DMCs are applied to the Poisson channel and finally Chapter 5 includes a plot of the results and some conclusions.
Chapter 2

Definitions and Notation

2.1 Notation

- $Q(x)$: Q-function, where $Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$
- $Q^{-1}(x)$: Inverse Q-function
- $\log x$ : is the natural logarithm such that $\log e = 1$
- $\epsilon$: Average probability of error

2.2 Channel Quantities

The following quantities are used frequently throughout this report. For a channel $W$, an input distribution $P$ and an output distribution $P_Y(y) = \sum_x W(y|x) P(x)$, we define:

- information density $i(x; y) = \log \frac{W(y|x)}{P_Y(y)}$;
- mutual information $I(P, W) = \mathbb{E}[i(X; Y)]$;
- unconditional absolute third order moment
  \[ T_u(P, W) = \mathbb{E}[|i(X; Y) - I(P, W)|^3]; \quad (2.1) \]
- unconditional information variance
  \[ U(P, W) = \text{Var}[i(X; Y)]; \quad (2.2) \]
• and the following quantity, which is defined in [3, Section 3.4.5], and sometimes referred to as the *reverse conditional information variance* or *reverse dispersion* [4]

\[
V^r(P, W) = \mathbb{E}[\text{Var}[i(X; Y)|Y]] = \sum_{x,y} P(x) W(y|x) \left[ \log^2 \frac{W(y|x)}{P_Y(y)} - \left( \mathbb{E}_{P_X|Y=y} \left[ \log \frac{W(y|X)}{P_Y(y)} \right] \right)^2 \right],
\]

The necessary and sufficient condition under which \( V^r(P, W) = 0 \), as explained in [3, Lemma 52], is that for all \( x \), such that \( P(x) > 0 \)

\[
W(y|x) \in \{h(y), 0\},
\]

for some function \( h(y) \). In other words, \( \forall x : P(x) > 0, W(y|x) \) is either 0 or some value that does not depend on \( x \).

### 2.3 Other Definitions

**Definition 2.1.** Consider two discrete channels described by the transition probabilities \( P_{Y_1|X} \) and \( Q_{Y_2|X} \). Channel \( Q \) is said to be *stochastically degraded* with respect to channel \( P \) if there exists a stochastic transformation \( Q'_{Y_2|Y_1} \) such that

\[
Q_{Y_2|X}(y_2|x) = \sum_{y_1} Q'_{Y_2|Y_1}(y_2|y_1) P_{Y_1|X}(y_1|x).
\]
Chapter 3

Upper Bound

Using the strategy mentioned in Section 1.4.1, upper bounds on the number of codewords $M$ that can be sent over a channel can be derived from lower bounds on the probability of error. A lower bound on the average probability of the binary erasure channel is already available in [3, Theorem 43]. Thus, if one could relate the performance of the BEC to the Poisson channel, the road to obtaining bounds on $\log M$ for the Poisson channel becomes a much smoother ride.

3.1 Lower Bounding Poisson Channel Error Probability

Proposition 3.1. The average error probability of the binary-input binary-output Poisson channel model described by (1.1) is lower-bounded by that of the binary erasure channel (3.1) for sufficiently small $\Delta$.

$$W_{\text{BEC}}(y|x) = \begin{bmatrix} k\Delta & 0 \\ 1 - k\Delta & 1 - k\Delta \\ 0 & k\Delta \end{bmatrix}, \quad (3.1)$$

$$\epsilon_{\text{avg}, \text{Poisson}(1.1)} \geq \epsilon_{\text{avg}, \text{BEC}(3.1)}. \quad (3.2)$$

To prove Proposition 3.1, we make use of Lemma 3.2.

Lemma 3.2. The binary-input binary-output Poisson channel model described by (1.1) is stochastically degraded (see Definition 2.1) with respect to
Chapter 3  Upper Bound

the binary erasure channel described by (3.1) for all $k$ such that

$$k \geq \max \left\{ Ase^{-A\Delta s}, A(s + 1)e^{-A\Delta(s+1)} \right\}.$$  \hspace{1cm} (3.3)

Proof. Consider the transformation matrix

$$Q' = \begin{bmatrix}
1 - \frac{Ase^{-A\Delta s}}{k} & 1 - \frac{A(s+1)e^{-A\Delta(s+1)}}{k} \\
\frac{Ase^{-A\Delta s}}{k} & 0
\end{bmatrix}.$$  \hspace{1cm} (3.4)

As long as we choose $k$ as some constant satisfying (3.3), the matrix $Q'$ is a stochastic matrix satisfying Definition 2.1. \hfill \square

Proof of Proposition 3.1. The proof is an application of the result in [5, Theorem 4.3.2], which states that if the following holds:

- $Q_2$ is a stochastically degraded channel with respect to some channel $Q_1$,
- $P$ is some distribution over $\mathcal{X}$,
- and $c(x, \gamma(y))$ is a measurable and bounded cost function,

where $\gamma(y)$ is a function signifying some policy, which we take here to be a decoding function, then we have the following result

$$J(P, Q_2) = \inf_{\gamma \in \Gamma} \int c(x, \gamma(y)) PQ_2(dx, dy) \geq \inf_{\gamma \in \Gamma} \int c(x, \gamma(y)) PQ_1(dx, dy) = J(P, Q_1).$$

Consider the encoder function

$$f : \{1, 2, \ldots, M\} \to \mathcal{X}^n,$$

the decoder function

$$g : \mathcal{Y}^n \to \{1, 2, \ldots, M\},$$

and the cost function

$$c(x^n, g(y^n)) = \mathbb{1}_{f^{-1}(x^n) \neq g(y^n)}.$$
For some discrete memoryless channel $W(y|x)$ and a distribution $P$ over $X^n$, define

$$J(P, W) = \inf_{g \in \mathcal{G}} \sum_{x^n, y^n} [P(x^n)c(x^n, g(y^n))W(y^n|x^n)]$$

$$= \inf_{g \in \mathcal{G}} \sum_{x^n, y^n} \left[ P(x^n)c(x^n, g(y^n)) \prod_{k=1}^{n} W(y_k|x_k) \right].$$  \hfill (3.5)

If one chooses $P$ in (3.5) to be the uniform distribution over codewords, such that

$$P(x^n) = \begin{cases} \frac{1}{M} & x^n \text{ is a codeword}, \\ 0 & x^n \text{ is not a codeword}, \end{cases}$$

then $J(P, W)$ is equal to the error probability of a channel $W$ with an encoder $f$, averaged over codewords and optimized over decoder functions $g$, where the average probability of error can be stated as

$$\epsilon_{\text{avg}} = \sum_{x^n, y^n} \frac{1}{M} \mathbb{1}(f^{-1}(x^n) \neq g(y^n))W(y^n|x^n).$$

Let $f_p, f_b$ be the optimal encoder functions for the Poisson channel (1.1) and the BEC channel (3.1), respectively. Then, the average error probability that can be achieved over the BEC channel when used with the encoder function $f_p$ and the related probability distribution $P_{\text{Poisson}}$ over $X^n$ must be at least as large that achieved when the channel is used with its optimal encoder function $f_b$. Since the Poisson channel (1.1) is stochastically degraded with respect to the BEC channel (3.1), their $n$-fold products are stochastically degraded too. Thus, we can apply [5, Theorem 4.3.2] to get

$$J(P_{\text{BEC}}, W_{\text{BEC}})|_{f_b} \leq J(P_{\text{Poisson}}, W_{\text{BEC}})|_{f_b} \leq J(P_{\text{Poisson}}, W_p)|_{f_p},$$

which proves (3.2).

\hfill \Box

### 3.2 Upper Bound on $\log M$

At this point, we may use the BEC average probability of error lower bound of [3, Theorem 43], to lower-bound that of the $2 \times 2$ Poisson channel (1.1). Then, for any $\Delta$, the upper bound on $\log M$ derived in [3, Theorem 44] for the BEC (3.1) is valid for the Poisson channel (1.1). Setting $n = T/\Delta$, we
Chapter 3 Upper Bound

obtain

\[ \log M \leq Tk - \sqrt{Tk(1-k\Delta)}Q^{-1}\left( \epsilon + \frac{[(k\Delta)^2 + (1-k\Delta)^2] + \frac{1}{2}\sqrt{2\pi}}{\sqrt{Tk(1-k\Delta)}} \right). \]

Taking the limit as \( \Delta \to 0 \), we obtain the following bound

\[ \log M \leq Tk - \sqrt{TkQ^{-1}}\left( \epsilon + \frac{1 + \frac{1}{2}\sqrt{2\pi}}{\sqrt{Tk}} \right). \quad (3.6) \]

Since we are interested in the case where \( \Delta \to 0 \), we assume that there exists \( \Delta_0 \) such that \( \forall \Delta \leq \Delta_0 \)

\[ Ase^{-A\Delta s} \leq A(s + 1)e^{-A\Delta(s+1)}, \]

so that we can take \( k = A(s + 1)e^{-A\Delta(s+1)} \), which in the limit is \( k = A(s + 1) \).

One could also minimize (3.6) over \( k \), subject to the condition (3.3), to obtain a tighter form of the bound (3.6).

3.3 Region of Validity

Since the arguments of the \( Q^{-1} \) function must be in \([0, 1]\), we obtain the following range of values of \( T \) where the bound (3.6) is valid:

\[ T \geq \left( \frac{1 + \frac{1}{2}\sqrt{2\pi}}{1 - \epsilon} \right)^2 \frac{1}{k}. \quad (3.7) \]
Chapter 4

Achievability Bound à la
Polyanskiy

In this chapter, we first apply the bounds derived in [3] for discrete memoryless
channels to the channel model (1.1), assuming some nonzero finite pulse-
width $\Delta$. In this step, one can assume $V^r(P,W) > 0$ and $U(P,W) > 0$. Consequently, the bounds (4.1) and (4.19) are valid. Then, we take the limit of the resulting expressions as $\Delta \to 0$.

The log $M$ achievability bounds derived in [3] are an application of the strategy described in Section 1.4.1. The difference between these bounds stems from the fact that each one was derived using a different upper estimate on the probability of error. For a general DMC with $V^r(P,W) > 0$, the bound derived using the Random Coding Union Bound (RCU) is tighter than that derived using the Dependence Testing Bound (DT). However, for our channel (1.1) as we take $\Delta \to 0$, this does not turn out to be the case.

4.1 Achievability using the RCU Bound

In his thesis [3, Theorem 53], Polyanskiy derives the following achievability bound for any DMC with $V^r(P,W) > 0$ using the RCU bound

$$\log M(n, \epsilon) \geq nI(P,W) + \sqrt{nU(P,W)Q^{-1}(\epsilon)} + \frac{1}{2} \log n + O(1).$$

(4.1)

This bound is derived from the Random Coding Union Bound (RCU), which is an upper bound on error probability. Since we set $n = T/\Delta$, we need to recover the $O(1)$ terms in (4.1). Thus, in what follows we elaborate
Chapter 4  Achievability Bounds

the arguments used in the proof of the bound (4.1) in [3]. Note, however, that some of the arguments used here are made specific to the case of the binary-input binary-output Poisson channel, and may not be valid for other channels.

4.1.1 Proof Ingredients

The proof of (4.1) makes use of the following arguments

- The Random Coding Union (RCU)
- Chernoff Bound

[3, Lemma 20] Let $Z_1, \ldots, Z_n$ be independent random variables such that $\sigma^2 = \sum_j \text{Var}[Z_j] > 0$ and $T = \sum_{j=1}^n \mathbb{E}[|Z_j - \mathbb{E}[Z_j]|^3] < \infty$, then for any $A$

$$
\mathbb{E}\left[e^{-\sum_{j=1}^n Z_j \mathbb{1}\{\sum_{j=1}^n Z_j > A\}}\right] \leq 2\left(\log \frac{2}{\sqrt{2\pi}} + \frac{12T}{\sigma^2}\right) \frac{1}{\sigma} e^{-A}.
$$

(4.2)

- Berry Esseen Theorem: consider the same setting as in [3, Lemma 20]. Then for any $\lambda$

$$
\left| \mathbb{P}\left[ \frac{\sum_{j=1}^n (Z_j - \mathbb{E}[Z_j])}{\sigma} \leq \lambda \right] - Q(-\lambda) \right| \leq c \frac{T}{\sigma^{3/2}},
$$

(4.3)

where $c = 1$ in case of i.i.d. $Z_j$, and $c = 6$ otherwise.

Random Coding Union Bound

The RCU bound is the starting point of the proof, and it is given by

$$
\epsilon \leq \mathbb{E}_{P^{X^n, Y^n}} \min\left\{1, MP_{X^n} [\bar{I}_n \geq I_n | X^n, Y^n] \right\},
$$

where

$$
I_n = \sum_{k=1}^n i(X_k; Y_k)
$$

and

$$
\bar{I}_n = \sum_{k=1}^n i(\bar{X}_k; Y_k),
$$

where $\bar{X}$ is independent of $X, Y$, such that $P_{X, \bar{X} | Y} = P_X P_{X | Y}$. Note that $X$ and $\bar{X}$ are identically distributed. This bound is derived using a random
coding argument and an application of the union bound. We rewrite the RCU bound in the following, more suggestive, form:

\[ \epsilon \leq \sum_{x^n, y^n} P(x^n, y^n) \min \left\{ 1, M \bar{P}_{X^n} \left[ \bar{I}_n(X^n, y^n) \geq I_n(x^n, y^n) \mid x^n, y^n \right] \right\}, \]

where one can see that conditioned on \(X^n, Y^n\), \(I_n\) becomes a deterministic variable, and due to the independence of \(\bar{X}^n\) of \(X^n\), one may write

\[
P_{X^n} \left[ \bar{I}_n(X^n, y^n) \geq I_n(x^n, y^n) \mid x^n, y^n \right] = P_{X^n} \left[ I_n(X^n, y^n) \geq I_n(x^n, y^n) \mid y^n \right].
\]

Define

\[ g(s, y^n) = P_{X^n} \left[ I_n(X^n, y^n) \geq s \mid y^n \right], \]

for some deterministic \(s\). Next, an upper bound on \(g(s, y^n)\) can be obtained using Bayes rule as explained in [6]

\[ g(s, Y^n) = \mathbb{E} \left[ e^{-I_n \mathbb{1}_{\{I_n \geq s\}}} \mid Y^n \right]. \]

**Chernoff Bound**

Next, the space of events is split\(^1\) into an event \(\mathcal{F}^c\) whose probability of occurrence is to be upper-bounded in this section, and \(\mathcal{F}\) which we will deal with subsequently. Define the random variables

\[ S_{Y_k} = \text{Var}[i(X_k, Y_k) \mid Y_k], \]

whose expectation is

\[ \mathbb{E}[S_{Y_k}] = V^r(P, W), \]

and which are independent. We wish to upper-bound the probability of the event

\[ \mathcal{F}^c = \left\{ \sum_{k=1}^n S_{Y_k} \leq V^r(P, W) \right\}. \]

Applying the Chernoff bound, one obtains

\[
P \left[ \sum_{k=1}^n S_{Y_k} \leq \frac{V^r(P, W)}{2} \right] = P \left[ e^{-t \sum_{k=1}^n S_{Y_k}} \geq e^{-t \frac{V^r(P, W)}{2}} \right] \leq e^{t \frac{V^r(P, W)}{2}} \mathbb{E} \left[ e^{-t \sum_{k=1}^n S_{Y_k}} \right], \forall t > 0.
\]

\(^1\)This approach is used in [3], but the need for it is not entirely clear.
Choosing $t = 1$, and since $S_{Y_k}$ are i.i.d., we can apply the previous bound to our binary input-binary output Poisson channel as follows

$$P \left[ \sum_{k=1}^{n} S_{Y_k} \leq \frac{V^r(P, W)}{2} \right] \leq e^{\frac{V^r(P, W)}{2}} \prod_{k} \mathbb{E} \left[ e^{-S_{Y_k}} \right]$$

$$= e^{\frac{V^r(P, W)}{2}} \left( P_Y(0) e^{-V_c(0)} + P_Y(1) e^{-V_c(1)} \right)^n,$$

where $V_c(0) = \text{Var}[i(X; Y)|Y = 0]$, and $V_c(1) = \text{Var}[i(X; Y)|Y = 1]$. The expressions for $V_c(0)$ and $V_c(1)$ for the channel (1.1) can be found in the appendix, equations (A.1) and (A.2).

So we may conclude that

$$P \left[ \sum_{k=1}^{n} S_{Y_k} \leq \frac{V^r(P, W)}{2} \right] \leq e^{\frac{V^r(P, W)}{2}} e^{-nK_1}, \quad (4.4)$$

where

$$K_1 = \min \{ V_c(0), V_c(1) \}. \quad (4.4)$$

To show that $K_1 > 0$, one may note that given the original assumption that the DMC has $V^r(P, W) > 0$, the condition (2.3) cannot be true. However, for the quantity

$$\text{Var}[i(X; Y)|Y = y] = \sum_{x} P(x) W(y|x) \left[ \log \frac{W(y|x)}{P_Y(y)} - \left( \mathbb{E}_{P_X|Y=y} \left[ \log \frac{W(y|X)}{P_Y(y)} \right] \right)^2 \right]$$

to be equal to zero for any $y$, the same condition (2.3) must be true $\forall x$ such that $P(x) > 0$. Thus,

$$K_1 = \min_{y} \text{Var}[i(X; Y)|Y = y] > 0.$$

**Applying [3, Lemma 20]**

Now, we need to upper-bound $g(I_n, Y_n)$, given that the event $\mathcal{F}$ has occurred using (4.2).

$$g(s, Y^n) = \mathbb{E} \left[ e^{-I_n 1_{\{I_n > s\}}} | Y^n \right]$$

$$\leq 2 \left( \frac{\log 2}{\sqrt{2\pi}} + \frac{12 T_y}{\sigma_y^2} \right) \frac{1}{\sigma_y} e^{-s}, \quad (4.5)$$
where

\[ \sigma_y^2 = \sum_{k=1}^{n} \text{Var}[i(X_k; Y_k)|Y^n] \]

\[ = \sum_{k=1}^{n} \text{Var}[i(X_k; Y_k)|Y_k] \quad \text{by independence} \]

\[ \geq \sum_{k=1}^{n} \min_y \text{Var}[i(X; Y)|Y = y] = nK_1. \]

and

\[ T_y = \sum_{k=1}^{n} \mathbb{E} \left[ |i(X_k, Y_k) - \mathbb{E}[i(X_k, Y_k)]|^2 |Y^n \right] \]

\[ = \sum_{k=1}^{n} \mathbb{E} \left[ |i(X_k, Y_k) - \mathbb{E}[i(X_k, Y_k)]|^2 |Y_k \right] \quad \text{by independence}. \]

Next, we need to upper-bound \( T_y \). Making use of the fact that \( Y_k \) is binary, we have

\[ T_y \leq n \max \{ T_c^{(0)}, T_c^{(1)} \}, \]

where \( T_c^{(0)} \) and \( T_c^{(1)} \) are evaluated in the appendix, Equations (A.3) and (A.4).

Thus, we can finally upper-bound (4.5) as follows

\[ g(s, Y^n) \leq \frac{K_2 e^{-s}}{\sqrt{n}}, \quad (4.6) \]

where

\[ K_2 = \frac{2}{\sqrt{K_1}} \left[ \frac{\log 2}{\sqrt{2\pi}} + \frac{12 \max \{ T_c^{(0)}, T_c^{(1)} \}}{K_1} \right]. \]

### 4.1.2 Achievability Bound Proof

We now incorporate all the previous arguments to obtain, in more detailed steps, the proof derived in [3, Theorem 53] for the achievability bound, with all the \( O(1) \) terms determined.

\[ \epsilon \leq \mathbb{E}_{P_{X^n, Y^n}} \left[ \min \{ 1, Mg(I_n, Y_n) \} \right] \quad (4.7) \]

\[ = P[\mathcal{F}^c] \mathbb{E}[\min \{ 1, Mg(I_n, Y_n) \} |\mathcal{F}^c] \]

\[ + P[\mathcal{F}] \mathbb{E}[\min \{ 1, Mg(I_n, Y_n) \} |\mathcal{F}] \]

\[ \leq P[\mathcal{F}^c] + P[\mathcal{F}] \mathbb{E}[\min \{ 1, Mg(I_n, Y_n) \} |\mathcal{F}] \quad (4.8) \]
\[ \leq e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P[\mathcal{F}] \mathbb{E}[\min\{1, Mg(I_n, Y_n)\} | \mathcal{F}] \] (4.10)

\[ \leq e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P[\mathcal{F}] \mathbb{E}\left[ \min\left\{ 1, \frac{MK_2}{\sqrt{n}} e^{-I_n} \right\} | \mathcal{F} \right] \] (4.11)

\[ = e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P[\mathcal{F}] \mathbb{E}\left[ e^{-I_n} \frac{MK_2}{\sqrt{n}} 1_{\{e^{-I_n} \frac{MK_2}{\sqrt{n}} \leq 1\}} | \mathcal{F} \right] \] (4.12)

\[ = e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P[\mathcal{F}] P\left[ I_n \leq \log \frac{MK_2}{\sqrt{n}} \right] \] + \frac{MK_2}{\sqrt{n}} \mathbb{E}\left[ e^{-I_n} 1_{\{I_n > \log \frac{MK_2}{\sqrt{n}}\}} \right] \] (4.13)

\[ \leq e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P\left[ I_n \leq \log \frac{MK_2}{\sqrt{n}} \right] \] + \frac{MK_2}{\sqrt{n}} \mathbb{E}\left[ e^{-I_n} 1_{\{I_n > \log \frac{MK_2}{\sqrt{n}}\}} \right] \] (4.14)

\[ \leq e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + P\left[ I_n \leq \log \frac{MK_2}{\sqrt{n}} \right] \] + \frac{MK_2K_3}{n} e^{-\log \frac{MK_2}{\sqrt{n}}} \] (4.15)

\[ \leq e^{-\frac{\nu(P,W)}{2}} e^{-nK_1} + Q\left( nI(P, W) - \log \frac{MK_2}{\sqrt{n}} \right) \] + \frac{T_0}{\sqrt{nU^{3/2}}} + \frac{K_3}{\sqrt{n}} \] (4.16)

where (4.8) partitions the expectation over the occurrence of \( \mathcal{F} \) and \( \mathcal{F}^c \), (4.9) follows by upper-bounding \( \mathbb{E}[\min\{1, \ldots \}] \) by 1, and (4.10) follows from (4.4). Moreover, (4.11) follows from (4.6) and (4.12) follows by breaking down the minimum into two indicator functions; one indicating that the first argument is the minimum, and one for the other case. Step (4.14) follows by observing that for any event \( A \)

\[ P[A] = P[A|\mathcal{F}] P[\mathcal{F}] + P[A|\mathcal{F}^c] P[\mathcal{F}^c] \geq P[A|\mathcal{F}] P[\mathcal{F}], \]

and that for a positive random variable \( Z \)

\[ \mathbb{E}[Z] = \mathbb{E}[Z|\mathcal{F}] P[\mathcal{F}] + \mathbb{E}[Z|\mathcal{F}^c] P[\mathcal{F}^c] \geq \mathbb{E}[Z|\mathcal{F}] P[\mathcal{F}]. \]

By one more application of (4.2) to (4.14), we reach (4.15), where

\[ K_3 = \frac{2}{\sqrt{U(P, W)}} \left[ \log \frac{2}{\sqrt{2\pi}} + \frac{12T_0}{U(P, W)} \right], \]

where \( U(P, W) \) is defined in (2.2). The final step of this proof is an application of the Berry Esseen theorem (4.3) to (4.15) to obtain (4.16). Now, as explained
4.2 Achievability Bound Using Dependence Testing Bound

in [3], equating the right hand side of (4.16) to \( \epsilon \) and solving for \( \log M \) yields the following achievability bound

\[
\log M \geq \frac{nI(P, W)}{\Delta} - \sqrt{nU(P, W)Q^{-1}} \left( \epsilon - e^{-\frac{1}{2}K_1 - \frac{T_u}{\sqrt{nU^{3/2}}} \epsilon} \right) - \log K_2 + \frac{1}{2} \log n
\]

(4.17)

4.1.3 Achievability in Continuous Time as \( \Delta \to 0 \)

To extend the bound (4.17) to our channel model (1.1), we need to set \( n = T/\Delta \), and take the limit as \( \Delta \to 0 \).

\[
\log M \geq \frac{T - n}{\Delta} \log \left( \frac{2\pi}{U} \right) + \log \frac{2\pi}{U} \left( \epsilon - e^{-\frac{1}{2}K_1 - \frac{T_u}{\sqrt{nU^{3/2}}} \epsilon} \right) - \log K_2 + \frac{1}{2} \log n
\]

(4.19)

First, we need to evaluate some limits

\[
\hat{V}^r(P, W) = \lim_{\Delta \to 0} \frac{V^r(P, W)}{\Delta} = \frac{A q s (1 - q) (s + 1)}{q + s} \left( \log \frac{s + 1}{s} \right)^2,
\]

\[
\lim_{\Delta \to 0} \sqrt{\Delta K_3} = \frac{2}{U} \left[ \log \frac{2}{\sqrt{2\pi}} + \frac{12 T_u}{U^2} \right],
\]

where \( \tilde{U} \) and \( \tilde{T}_u \) are defined in (4.21) and (4.22), respectively. However,

\[
\lim_{\Delta \to 0} \log \left( \frac{2\pi}{U} \right) = -\infty,
\]

(4.18)

implying that the bound is loose.

4.2 Achievability Bound Using Dependence Testing Bound

Polyanskiy also derived a lower bound in [3, Theorem 47] for DMCs with \( U(P, W) > 0 \) using the Dependence Testing Bound [3, Theorem 18], which we use to obtain the following lower bound by setting \( n = \frac{T}{\Delta} \)

\[
\log M \geq \frac{T I(P, W)}{\Delta} - \sqrt{\frac{TU(P, W)}{\Delta}Q^{-1}} \left( \epsilon - \frac{2\log \frac{2\sqrt{\Delta}}{\sqrt{2\pi TU(P, W)}}}{\sqrt{T}} - \frac{5B}{\sqrt{T}} \right).
\]

(4.19)
where
\[ B = \frac{T_u(P, W)}{U(P, W)^2}, \]
and \( T_u(P, W) \) is the unconditional absolute third order moment as defined in (2.1). Letting \( \Delta \to 0 \),
\[ \tilde{I} = \lim_{\Delta \to 0} \frac{I(P, W)}{\Delta} = Aq \log\left(\frac{s+1}{q+s}\right)(s+1) - A s \log\left(\frac{s}{q+s}\right) (q-1), \quad (4.20) \]
\[ \tilde{U} = \lim_{\Delta \to 0} \frac{U(P, W)}{\Delta} = Aq \log^2\left(\frac{s+1}{q+s}\right)(s+1) - A s \log^2\left(\frac{s}{q+s}\right) (q-1), \quad (4.21) \]
\[ \tilde{T}_u = \lim_{\Delta \to 0} \frac{T_u(P, W)}{\Delta} = Aq(s+1)\log\left(\frac{s+1}{q+s}\right)^3 - A(q-1)s\log\left(\frac{s}{q+s}\right)^3, \quad (4.22) \]
thus, in the limit, we obtain the following bound
\[ \log M \geq \tilde{T} \hat{I} - \sqrt{\tilde{T} \hat{U} q^{-1}} \left( \epsilon - \frac{2 \log 2}{\sqrt{2 \pi T \epsilon}} - \frac{5 \tilde{T}_u}{\sqrt{T U^2}} \right). \]

### 4.2.1 Region of Validity

Since the arguments of the \( Q^{-1} \) function must be in [0, 1], we obtain the following range of values of \( T \) where the bound is valid:
\[ T \geq \left( \frac{5 \tilde{T}_u}{\hat{U}^2 \epsilon} + \frac{2 \log 2}{\sqrt{2 \pi U \epsilon}} \right)^2. \quad (4.23) \]
Chapter 5

Result and Conclusion

Figure 5.1 shows a plot of the previously obtained upper and lower bounds, where the start of the validity region of the lower bound is $T = 5.76 \times 10^{11}$, and for the upper bound it is $T = 0.0279$. While this report explored the extension of DMC bounds to the continuous-time Poisson channel, more work is still needed to find tighter bounds.

![Figure 5.1: Lower Bound obtained via (4.19) and Upper Bound by (3.6), A=100, s=0.25.](image)

Figure 5.1: Lower Bound obtained via (4.19) and Upper Bound by (3.6), A=100, s=0.25.

extension of DMC bounds to the continuous-time Poisson channel, more work is still needed to find tighter bounds.

19
Appendix A

A.1 Conditional Variance and Third Order Absolute Moment of Channel (1.1)

Let $w(0|0) = p_0$ and $w(1|1) = p_1$

$V_c^{(0)}$

\[
V_c^{(0)} = E \left[ (i(X_k, Y_k) - E[i(X_k, Y_k)])^2 | Y_k = 0 \right]
\]

\[
= p_0 (1 - q) \left[ \log \frac{p_0}{P_Y(0)} - \frac{p_0 (1 - q) \log \frac{P_Y(0)}{P_0}}{P_Y(0)} + \frac{q (1 - p_1) \log \frac{P_Y(0)}{1 - p_1}}{P_Y(0)} \right]^2
\]

$V_c^{(1)}$

\[
V_c^{(1)} = E \left[ (i(X_k, Y_k) - E[i(X_k, Y_k)])^2 | Y_k = 1 \right]
\]

\[
= p_1 q \left[ \log \frac{p_1}{P_Y(1)} - \frac{(1 - p_0) (1 - q) \log \frac{1 - P_Y(0)}{P_Y(1)}}{P_Y(1)} - \frac{p_1 q \log \frac{p_1}{P_Y(1)}}{P_Y(1)} \right]^2
\]

$T_c^{(0)}$

\[
T_c^{(0)} = E \left[ |i(X_k, Y_k) - E[i(X_k, Y_k)]|^3 | Y_k = 0 \right]
\]

\[
= p_0 (1 - q) \left| \log \frac{p_0}{P_Y(0)} \right|^3 + \frac{p_0 \log \frac{P_Y(0)}{P_0}}{P_Y(0)} (1 - q) + \frac{q \log \frac{P_Y(0)}{1 - p_1}}{P_Y(0)} (1 - p_1) \right]^3
\]
\[ T_c^{(1)}(1) = E \left[ i(X_k, Y_k) - E[i(X_k, Y_k)] \right] | Y_k = 1 \]

\[ = \frac{p_1 q}{p_Y(1)} \left| \log \frac{p_1}{p_Y(1)} + \frac{(1 - p_0)(1 - q) \log \frac{1 - p_0}{p_Y(1)}}{-p_Y(1)} + \frac{p_1 q \log \frac{p_1}{p_Y(1)}}{-p_Y(1)} \right|^3 \]
Bibliography


List of Figures

5.1 Lower Bound obtained via (4.19) and Upper Bound by (3.6), $A=100, s=0.25$.

19