(4)

Capacity of the Poisson Channel with Time-Varying Noise Intensity and Jamming

Michael R. Frey

Department of Statistics
University of North Carolina
Chapel Hill, NC 27599

May 1988
LISS 25

Abstract - The information capacity of the Poisson channel with jamming and nonrandom time-varying noise intensity is obtained for a time-varying peak constraint on the encoder intensity. The optimal jamming signal is determined. Feedback is shown not to increase the channel capacity for the case of nonrandom noise intensity. Poisson channels with thinning are introduced.

Supported by ARO Contract DAAL03-86-G-0022 and ONR Contract N00014-86-K-0039.

These results were presented in part at the 22nd Annual Conference on Information Sciences and Systems, March 16-18, 1988, Princeton University.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Department of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0039
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: NR PROJECT NO.: 042 TASK NO.: 269 WORK UNIT NO.: SRO 105
11. TITLE (Include Security Classification) Capacity of the Poisson Channel ...	

12. PERSONAL AUTHOR(S)	M.R. Frey

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
TECHNICAL	FROM ____ TO ____	May 1988	25

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Information capacity; Poisson channel; Jamming; Mutual information; Thinning.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The information capacity of the Poisson channel with jamming and nonrandom time-varying noise intensity is obtained for a time-varying peak constraint on the encoder intensity. The optimal jamming signal is determined. Feedback is shown not to increase the channel capacity for the case of nonrandom noise intensity. Poisson channels with thinning are introduced.

11. TITLE CONT.: With Time-Varying Noise Intensity and Jamming

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

I. INTRODUCTION

Optical communications systems modeled by a Poisson channel (Fig. 1) are described in [6], [10], [11], [12], [13], and [15]. The noise intensity (shown) in the channel model represents, nominally, noise due to background radiation as seen by the receiver. In this role, modeling the noise intensity by a real constant may not be adequate. For example, models in which the noise (-plus-signal) intensity is the magnitude of a time-varying Gaussian vector appear in the literature [8], [16].

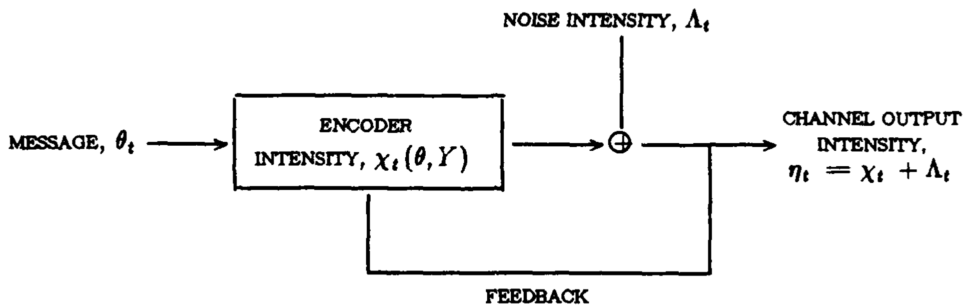

Figure 1. Poisson channel model.

Besides representing background radiation, the noise intensity in the channel model may be used, also, to represent a variety of other features of the modeled optical system. Background radiation will generally have a much wider spectrum than that of the channel source. Thus, the optical filter used in the receiver to reduce the apparent background radiation may be factored into the noise intensity. The dark current generated in the receiver photon detector adds to its output current and is another source of noise. This noise, when referred back to the receiver input, contributes to the channel noise intensity. The receiver may be illuminated by a jammer - another noise source. Also, imperfect encoder modulation (where the encoder is never completely off or intersymbol interference is present) in effect contributes to the noise.

For many optical systems, the encoder in the channel model is interpreted as modulating the envelope of the output of a photon source. The photon detector receiving the stream of photons operates with a certain efficiency $q \leq 1$; meaning an incident photon generates an electron-hole carrier pair with probability q. Three types of photon detectors are commonly considered in direct detection optical receivers [4]: p-i-n photodiodes, avalanche photodiodes, and photoconductors. Most present fiber optic systems use the p-i-n photodiode; however, the internal gain of the avalanche photodiode makes it an increasingly attractive choice. Carrier pairs generated in the avalanche photodiode go on to generate additional pairs by a process of collision ionization depending upon the type of carrier (hole or electron), the semiconductor material out of which the device is constructed, and the voltage across the device. The current generated by collision ionization has a message dependent noise component.

Background radiation, optical filtering, dark current, jamming, imperfect encoder modulation, conversion efficiency, and collision ionization - because of these and other factors, the engineer may, in the design and analysis of optical communications systems, endow the channel noise intensity with a rather rich structure. Thus a complete theory of channel capacity for the Poisson channel should address time-varying, random noise intensities dependent on the message process θ. In this paper we take several steps in this direction and give channel capacity results for time-varying noise intensity and time-varying encoder constraints. Results for jamming are also given. To begin, we specify the channel model and recall some definitions from information theory.

The Poisson channel model addressed in this paper is represented in Figure 1. The channel output $Y=\{Y_t\}_{0\le t\le T}$ is a Poisson-type (simple) point process directed by the stochastic intensity $\eta_t = \chi_t + \Lambda_t$. Y and the message process $\theta=\{\theta_t\}_{0\le t\le T}$ are defined on a common probability space $(\Omega,\boldsymbol{F},P)$ with respective completed natural histories $\boldsymbol{F}_t^\theta$ and $\boldsymbol{F}_t^Y$. The message θ is encoded into the channel encoder output via the encoder intensity χ_t. $\chi_t = \chi_t(\theta,Y)$ is required to be an $\boldsymbol{F}_t^\theta \vee \boldsymbol{F}_t^Y$-adapted encoding of the message process θ and channel output Y permitting causal message encoding and noiseless, nonanticipative, instantaneous feedback. The channel noise process N_t is the Poisson-type point process directed by the intensity Λ_t. We take the processes θ and N to be independent and impose, on the encoder intensity, a time-varying peak constraint

$$0 \le \chi_t \le c(t) \tag{1}$$

for all $t \in [0,T]$ where $c(t)$ is nonnegative, bounded, and Lebesgue-measurable or an average constraint

$$E\left[\int_0^T \chi_t\, dt\right] \le k_0 T \qquad (2)$$

for some positive constant k_0. We treat the case in which the noise intensity is time-varying and deterministic: $\Lambda_t = \lambda(t)$. Further discussion of the Poisson channel model can be found in [1], [3], [7], and [15] and the references cited therein.

To define channel capacity let μ_θ, μ_Y, and $\mu_{\theta Y}$ be the marginal and joint measures induced by the message and output processes, θ and Y, on the spaces S_θ, S_Y, and $S_\theta \times S_Y$ where S_θ and S_Y are the spaces of trajectories of θ and Y on $[0,T]$. Write the induced product measure as $\mu_{\theta\times Y}$. Then, the ***average mutual information*** in θ and Y over the interval $[0,T]$ is

$$I^T[\theta,Y] = E\left[\ln \frac{d\,\mu_{\theta Y}}{d\,\mu_{\theta\times Y}}\right]$$

provided $\mu_{\theta Y} \ll \mu_{\theta\times Y}$; otherwise $I^T[\theta,Y]=\infty$. The channel ***information capacity*** is

$$C = \sup_\theta \sup_\chi \frac{1}{T} I^T[\theta,Y]$$

where θ is any jointly measurable process on $[0,T]$ and $\chi_t = \chi_t(\theta,Y)$ is any $\boldsymbol{F}_t^\theta \vee \boldsymbol{F}_t^Y$-adapted mapping. For the Poisson channel more revealing expressions for the average mutual information are [Theorem 19.3, 9]:

$$I^T[\theta,Y] = E\left[\int_0^T (\eta_t \ln\eta_t - \eta_t \ln\hat{\eta}_t)\,dt\right] \qquad (3)$$

and

$$I^T[\theta,Y] = E\left[\int_0^T (\eta_t \ln\eta_t - \hat{\eta}_t \ln\hat{\eta}_t)\,dt\right] \qquad (4)$$

where $\hat{\eta}_t = E[\eta_t \mid \boldsymbol{F}_t^Y]$. These expressions are valid for nonrandom noise intensities.

II. CHANNEL CAPACITY FOR TIME-VARYING NONRANDOM NOISE INTENSITY

The capacity of the Poisson channel is known for the case in which the noise intensity is a real constant, $\Lambda_t = \lambda$. For nontime-varying peak-constrained encoder intensity $c(t) = c$ (no average constraint imposed), Kabanov [7] showed the channel capacity to be

$$\mathbf{C} = C(\lambda, c) \tag{5}$$

where

$$C(x,y) = \frac{x}{e}\left(1 + \frac{y}{x}\right)^{1+x/y} - x\left(1 + \frac{x}{y}\right)\ln\left(1 + \frac{y}{x}\right). \tag{6}$$

We show by way of Lemmas 1, 2, and 3 and Theorem 4 that the approach taken by Kabanov can be adapted to extend his result to the case in which the noise intensity is nonrandom and time-varying: $\Lambda_t = \lambda(t)$ and in which the peak constraint is also time-varying. Lemma 1 treats the case in which $\lambda(t)$, $c(t)$ are simple functions. Lemma 2 goes on to treat extended real-valued simple functions $\lambda(t)$. Then in Theorem 4, for time-varying noise intensity $\lambda(t)$ and time-varying peak-constrained encoder output, we obtain the general result

$$\mathbf{C} = \frac{1}{T}\int_0^T C(\lambda(t), c(t))dt . \tag{7}$$

Note that when the noise intensity and peak constraint are constant, $\lambda(t) = \lambda$ and $c(t) = c$, (7) gives $\mathbf{C} = C(\lambda, c)$ in accordance with (5).

In the proof of Lemma 1 a sequence of message processes $\theta^{(m)}$ and an encoding scheme χ are introduced which give average mutual information in the channel arbitrarily close, as $m \to \infty$, to the channel capacity. The encoding scheme there does not use feedback. So, for nonrandom time-varying noise intensity, feedback does not increase channel capacity, i.e., the average mutual information in the channel can be made arbitrarily close to the channel capacity without use of feedback if the message process θ and the encoding scheme χ are appropriately chosen. This is so when the encoder intensity is peak-constrained, average-constrained, or both.

Davis [3] extended Kabanov's result (5) to reflect the imposition of an average constraint on the encoder output. Davis' reasoning applies as well to the case of time-varying noise intensity.

We find that, for nonrandom, time-varying noise intensity $\lambda(t)$ and an encoder intensity both peak- and average-constrained as per (1) and (2) with $c(t) = c$, the channel capacity is

$$C = \begin{cases} \frac{1}{T}\int_0^T \left[\frac{\phi(c,\lambda(t))}{c} k_0 - \phi(k_0,\lambda(t)) \right] dt \ , \text{ for } k_0 \le \frac{c}{e} \\ \frac{1}{T}\int_0^T C(\lambda(t) \wedge \lambda_0, c)\, dt \ , \text{ for } \frac{c}{e} < k_0 < \frac{c}{2} \\ \frac{1}{T}\int_0^T C(\lambda(t), c)\, dt \ , \text{ for } k_0 \ge \frac{c}{2} \end{cases} \tag{8}$$

where $\lambda(t) \wedge \lambda_0$ is the minimum of $\lambda(t)$ and λ_0,

$$\phi(x,y) = (x+y)\ln(x+y) - y\ln y \ ,$$

and λ_0 is uniquely defined, for $\frac{c}{e} < k_0 < \frac{c}{2}$, by

$$k_0 = \frac{\lambda_0}{e}\left(1 + \frac{c}{\lambda_0}\right)^{1+\lambda_0/c} - \lambda_0 \, .$$

Of course, the average constraint is inoperative when $k_0 \ge c$; now we see from (8) that the average constraint is inoperative even when $k_0 \ge c/2$. (8) adjusts in the obvious way to accommodate a time-varying peak constraint $c(t)$.

We close this section with one final result for the capacity of the Poisson channel. We still suppose that the encoder intensity is peak-constrained $0 \le \chi_t \le c(t)$ but no longer take $c(t)$ to be a given function. Instead, we suppose $c(t)$ may be chosen freely subject only to the constraint

$$\frac{1}{T}\int_0^T c(t)\, dt \le P$$

for some given $P > 0$. Then for a known nonrandom noise intensity $\Lambda_t = \lambda(t)$ the channel capacity is found to be $C = P/e$. The proof of this result (Theorem 6) given in Section V recognizes the need to concentrate the power $c(t)$ available to the encoder into as short a time interval as possible. Thus one obtains average mutual information in the channel closer and closer to P/e by choosing

$$c(t) = \frac{P}{|A|} 1_A(t)$$

and letting $| A | \rightarrow 0$. This is reminiscent of Davis' result [3] for polarization modulation. Davis showed that channel capacity is maximized when encoder power is not distributed over two orthogonally polarized, separately modulated channels but, instead, is concentrated solely in one channel. This was because of the convexity of $C(x, y)$ in y. For this same reason it is also better not to distribute encoder power over time but, rather, to concentrate it into as short a time interval as possible. Also, concentration of the encoder intensity into a short interval permits it to be very large; swamping out whatever noise intensity may be present in that interval. Hence the result that $\lambda(t)$ does not appear in the expression for the capacity $C = P / e$.

III. JAMMING

In this section we present results for the Poisson channel subjected to jamming. We make the presence of jamming explicit in the channel model by taking the channel intensity to be $\eta_t = \chi_t + \Lambda_t + J(t)$ where $J(t)$, the jamming intensity, is known, deterministic, nonnegative, and Lebesgue-measurable. We do not consider stochastic jamming intensities. Restricting attention to channels with nonrandom noise intensity, $\Lambda_t = \lambda(t)$, we have

$$\eta_t = \chi_t + \lambda(t) + J(t). \tag{9}$$

Then by Theorem 4, for a peak-constrained encoder, the channel capacity is

$$\mathbf{C} = \frac{1}{T}\int_0^T C(\lambda(t) + J(t), c(t))dt .$$

(If the encoder is also average-constrained we have the result analogous to (8)).

In general, the jammer seeks to minimize the channel capacity subject to constraints on the jamming intensity. If the jamming intensity is peak-constrained, $J(t) \le c_J$ for some $c_J > 0$, the solution is immediate. The optimal choice of jamming intensity is

$$J_{opt}(t) = c_J \quad \text{a.e.}$$

(because C(x,y) is positive and strictly decreasing in x). A more interesting problem arises when the jamming intensity is average-constrained,

$$\int_0^T J(t)dt \le PT .$$

For this case (Theorem 5),

$$J_{opt}(t) = [\alpha - \lambda(t)]^+ \quad \text{a.e.}$$

where α is determined from

$$\int_0^T [\alpha - \lambda(t)]^+ dt = PT \tag{10}$$

and x^+ is the positive part of $x \in \mathbf{R}$. Of course, other constraints or combinations of constraints may be imposed on the jamming intensity, leading to different forms for the optimal jamming intensity.

IV. THINNING

Signals transmitted through a Poisson channel can be viewed as a train of point events. The Poisson channel noise process corrupts the signal by interjecting extra points into the train of signal points. One can just as well imagine the signal to be corrupted by removal of points. This kind of signal corruption is called *thinning*. Thinning can be used to model an absorptive channel medium, optical filters in the receiver, or conversion efficiency of the receiver detector. Thinning may operate alone in a point process channel or in concert with some form of Poisson noise. The case in which points are removed (or not) independent of the removal of other points is called *independent thinning.*

Definition: Let C_t be a point process with

$$C_t = \sum_{k=0}^{\infty} 1_{[T_k,\infty)}(t)$$

where $\{T_k\}$ is a sequence of stopping times. Let $\{\beta_k\}$ be a sequence of i.i.d. Bernoulli r.v.s with $P\{\beta_k = 1\} = p(T_k)$. Then

$$\tilde{C}_t = \sum_{k=0}^{\infty} \beta_k 1_{[T_k,\infty)}(t)$$

is a probability-$p(t)$ independently thinned point process.

While many different types of thinning are conceivable, independent thinning is distinguished by being simple in concept, by being motivated by physical considerations and by easily admitting results for channel capacity. In particular, probability-$p(t)$ independent thinning of a Poisson process with intensity $i(t)$ produces a Poisson process with intensity $p(t)i(t)$ [2].

Consider the Poisson channel with noise intensity $\Lambda_t = \lambda(t)$ for $0 \le t \le T$ and suppose the noise is probability-$p(t)$ independently thinned. Let the encoder intensity χ_t be $\mathbf{F}_t^Y \vee \mathbf{F}_t^\theta$

-adapted with peak constraint $0 \leq \chi_t \leq c(t)$ and suppose the encoder output is probability-$q(t)$ independently thinned. Finally, suppose the thinnings of the encoder output and the noise are independent. A probability-$q(t)$ independently thinned encoder output with peak-constrained intensity $0 \leq \chi_t \leq c(t)$ is equivalent to an unthinned encoder output with intensity constraint $0 \leq \chi_t \leq c(t)q(t)$. Thus the channel capacity is

$$\mathbf{C} = \frac{1}{T}\int_0^T C(p(t)\lambda(t), q(t)c(t))dt \ .$$

For $p(t)=q(t)=p_o$, thinning reduces the capacity by a factor of p_o:

$$\mathbf{C} = \frac{p_o}{T}\int_0^T C(\lambda(t), c(t))dt \ .$$

(Here we have used $C(xz, yz) = zC(x, y)$.)

V. Results and Proofs

Lemma 1: Let the Poisson-type point process $Y = \{Y_t\}_{0 \le t \le T}$ directed by the intensity $\eta_t = \tilde{\lambda}(t) + \chi_t$ be the output of the Poisson channel with noise intensity given by the simple function

$$\tilde{\lambda}(t) = \sum_{j=1}^{n} \lambda_j 1_{E_j}(t), \quad t \in [0,T],$$

with $0 \le \lambda_j < \infty$ for all j and such that $\{E_j, 1 \le j \le n < \infty\}$, is a Lebesgue-measurable partition of $[0,T]$ (i.e. the E_j are Lebesgue-measurable subsets of $[0,T]$ and the union of the E_j is $[0,T]$.) Constrain the encoder $\chi_t = \chi_t(\theta, Y)$ to be an $\boldsymbol{F}_t^{\theta} \vee \boldsymbol{F}_t^{Y}$-adapted function of Y and the message process $\theta = \{\theta_t\}_{0 \le t \le T}$. Further require that $0 \le \chi_t \le \tilde{c}(t)$ where $\tilde{c}(t)$ is the simple function

$$\tilde{c}(t) = \sum_{i=1}^{l} c_i 1_{F_i}(t), \quad t \in [0,T]$$

where $\{F_i, 1 \le i \le l < \infty\}$ is a Lebesgue-measurable partition of $[0,T]$ and $0 \le c_i < \infty$ for all i. Then

$$\mathbb{C} = \frac{1}{T}\int_0^T C(\tilde{\lambda}(t), \tilde{c}(t))dt .$$

Proof: First we show

$$\mathbb{C} \le \frac{1}{T}\int_0^T C(\tilde{\lambda}(t), \tilde{c}(t))dt .$$

$\eta_t = \tilde{\lambda}(t) + \chi_t$ is bounded so by (4), applying Fubini's theorem,

$$I^T[\theta, Y] = \int_0^T \Big[E[\eta_t \ln \eta_t] - E[\hat{\eta}_t \ln \hat{\eta}_t] \Big] dt$$

$$= \int_0^T \Big[E[\phi(\chi_t, \tilde{\lambda}(t))] - E[\phi(\hat{\chi}_t, \tilde{\lambda}(t))] \Big] dt$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{l} \int_{E_j \cap F_i} \Big[E[\phi(\chi_t, \lambda_j)] - E[\phi(\hat{\chi}_t, \lambda_j)] \Big] dt$$

since $\tilde{\lambda}(t) = \lambda_j$ for $t \in E_j$. $\frac{\partial^2 \phi(x,y)}{\partial x^2} = \frac{1}{x+y} \geq 0$ for $x, y \geq 0$ so, by Jensen's inequality,

$$E[\phi(\hat{\chi}_t, \lambda_j)] \geq \phi(E[\hat{\chi}_t], \lambda_j) .$$

Using $E[\hat{\chi}_t] = E[\chi_t]$ and $\phi_j(x) \equiv \phi(x, \lambda_j)$ we get

$$I^T[\theta, Y] \leq \sum_{j=1}^{n} \sum_{i=1}^{l} \int_{E_j \cap F_i} \Big[E[\phi_j(\chi_t)] - \phi_j(E[\chi_t])\Big] dt .$$

We adopt a convenient notation: for a real measurable function, g,

$$E[g(\chi_t)] = \int_{\Omega} g(\chi_t(\omega)) P(d\omega) = \int_0^{\bar{c}(t)} g(x) P \circ \chi_t^{-1}(dx) \equiv Q(g) .$$

Q is indexed by the time t, or equivalently, by the probability measure $P \circ \chi_t^{-1}$. In terms of Q,

$$I^T[\theta, Y] \leq \sum_{j=1}^{n} \sum_{i=1}^{l} \int_{E_j \cap F_i} [Q(\phi_j) - \phi_j(Q(\mathrm{i}))] dt$$

where i is the identity function $\mathrm{i}(x) = x$. Let D_i be the set of all probability measures on $([0,c_i], \mathcal{B}([0,c_i]))$. Then

$$I^T[\theta, Y] \leq \sum_{j=1}^{n} \sum_{i=1}^{l} | E_j \cap F_i | \sup_{m \in D_i} \Big[Q(\phi_j) - \phi_j(Q(\mathrm{i}))\Big]$$

where $| E_j \cap F_i |$ is the Lebesgue measure of the set $E_j \cap F_i$. Therefore

$$\mathcal{C} \leq \frac{1}{T} \sum_{j=1}^{n} \sum_{i=1}^{l} | E_j \cap F_i | \sup_{m \in D_i} \Big[Q(\phi_j) - \phi_j(Q(\mathrm{i}))\Big] . \tag{11}$$

For the details of the calculation of the supremum on the right-hand side of (11) we refer to p.182 of [1]. There the supremum is shown to be

$$\sup_{m \in D_i} \Big[Q(\phi_j) - \phi_j(Q(\mathrm{i}))\Big] = C(\lambda_j, c_i)$$

where $C(x,c) = C(x)$ given in (6). Hence

$$\mathcal{C} \leq \frac{1}{T} \sum_{j=1}^{n} \sum_{i=1}^{l} | E_j \cap F_i | C(\lambda_j) = \frac{1}{T} \int_0^T C(\tilde{\lambda}(t), \bar{c}(t)) dt .$$

Now we show

$$C \geq \frac{1}{T}\int_0^T C(\tilde{\lambda}(t), \tilde{c}(t))dt \ .$$

Our approach is to find a sequence of message processes $\{\theta^{(m)}\}$ and an encoding χ such that

$$\lim_{m \to \infty} \frac{1}{T} I^T[\theta^{(m)}, Y] = \frac{1}{T}\int_0^T C(\tilde{\lambda}(t), \tilde{c}(t))dt \ . \tag{12}$$

To construct $\{\theta^{(m)}\}$ first define a sequence $\{\vec{V}^{(m)}\}$ of $l \times n$-dimensional matrices of independent telegraph signals $V_{ij,t}^{(m)}$, $i = 1,2,\ldots,l$, $j = 1,2,\ldots,n$, i.e., right-continuous homogeneous Markov jump processes with states $\{0,1\}$ (Chapter 3 of [5]). We choose the processes $V_{ij,t}^{(m)}$, $m \geq 1$, to have infinitesimal parameter matrices

$$Q^{(m)} = \begin{bmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{bmatrix} = \begin{bmatrix} -m & m \\ m(1-p)/p & -m(1-p)/p \end{bmatrix}$$

where $p = k_j / c_i$ with

$$k_j = \frac{\lambda_j}{e}\left(1 + \frac{c_i}{\lambda_j}\right)^{1+\lambda_j/c_i} - \lambda_j \ .$$

Also, for $m \geq 1$, let $P\{V_{ij,0}^{(m)} = 1\} = p$ so that $V_{ij,t}^{(m)}$ is stationary. Then define $\theta^{(m)}$ for all $m \geq 1$ as

$$\theta_t^{(m)} = \sum_{j=1}^{n}\sum_{i=1}^{l} V_{ij,\tau(t)}^{(m)} 1_{E_j \cap F_i}(t)$$

where

$$\tau(t) = \int_0^t 1_{E_j \cap F_i}(s)ds \ .$$

Also, let $\chi_t(\theta, Y) = \tilde{c}(t)\theta_t$. Then

$$I^T[\theta^{(m)}, Y] = \int_0^T \Big[E[\phi(\tilde{c}(t)\theta_t^{(m)}, \tilde{\lambda}(t))] - E[\phi(\tilde{c}(t)\hat{\theta}_t^{(m)}, \tilde{\lambda}(t))]\Big] dt$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{l} \int_{E_j \cap F_i} \Big[E[\phi(c_i V_{ij,\tau(t)}^{(m)}, \lambda_j)] - E[\phi(c_i \hat{V}_{ij,\tau(t)}^{(m)}, \lambda_j)]\Big] dt$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{l} \int_0^{|E_j \cap F_i|} \Big[E[\phi_j(c_i V_{ij,\tau}^{(m)})] - E[\phi_j(c_i \hat{V}_{ij,\tau}^{(m)})]\Big] d\tau \ .$$

Now

$$E[\phi_j(c_i V_{ij,t}^{(m)})] = \phi_j(0)P\{V_{ij,t}^{(m)}=0\} + \phi_j(c_i)P\{V_{ij,t}^{(m)}=1\} = p\,\phi_j(c_i)$$

since $\phi_j(0) = 0$ and $P\{V_{ij,t}^{(m)}=1\} = p$. $p\,\phi_j(c_i) - \phi_j(pc_i) = C(\lambda_j, c_i)$ so,

$$I^T[\theta, Y] = \sum_{j=1}^{n} \sum_{i=1}^{l} \int_0^{|E_j \cap F_i|} \left[C(\lambda_j, c_i) + \phi_j(pc_i) - E[\phi_j(c_i V_{ij,\tau}^{(m)})] \right] d\tau .$$

Thus to show (12) and complete the proof, it only remains to show

$$\lim_{m\to\infty} E[\phi_j(c_i \hat{V}_{ij,t}^{(m)})] = \phi_j(pc_i) .$$

The derivative of ϕ_j is bounded on $[0, c_i]$, say by a constant M, so $|\phi_j(c_i \hat{V}_{ij,t}^{(m)}) - \phi_j(pc_i)| \le Mc_i\, |\hat{V}_{ij,t}^{(m)} - p|$. Then

$$\begin{aligned}
|E[\phi_j(c_i \hat{V}_{ij,t}^{(m)})] - \phi_j(pc_i)| &\le E\left[|\phi_j(c_i \hat{V}_{ij,t}^{(m)}) - \phi_j(pc_i)| \right] \\
&\le Mc_i E\left[|\hat{V}_{ij,t}^{(m)} - p| \right] \\
&\le Mc_i \left[E[|\hat{V}_{ij,t}^{(m)} - p|^2] \right]^{1/2} \\
&\le Mc_i \left[E[\hat{V}_{ij,t}^{(m)\,2}] - p^2 \right]^{1/2} .
\end{aligned}$$

Beginning on p. 184 of [1] it is shown that $\lim_{m\to\infty} E[\hat{V}_{ij,t}^{(m)\,2}] = p^2$ so the proof is completed.

Lemma 2: $C(x, c) \to 0$ as $x \to \infty$ so $C(\infty, c) \equiv 0$. Then, with the encoder intensity peak-constrained by the nonnegative simple function $\tilde{c}(t)$,

$$\mathcal{C} = \frac{1}{T}\int_0^T C(\tilde{\lambda}(t), \tilde{c}(t))\,dt$$

for all nonnegative extended real-valued simple functions

$$\tilde{\lambda}(t) = \sum_{j=1}^{n} \lambda_j 1_{E_j}(t)$$

where $0 \le \lambda_j \le \infty$.

Proof: We need only consider

$$\tilde{\lambda}(t) = \sum_{j=1}^{n} \lambda_j 1_{E_j}(t)$$

where $\lambda_n = \infty$ and $0 \le \lambda_j < \infty$ for $j < n$. Define

$$f(x;Y) = (x+Y)\ln\left(\frac{x+Y}{x+\hat{Y}}\right) .$$

Then, using (3),

$$I^T[\theta,Y] = \int_0^T E[f(\tilde{\lambda}(t);\chi_t)]dt$$

$$= \sum_{j=1}^{n} \int_{E_j} E[f(\lambda_j;\chi_t)]dt$$

$$= \sum_{j=1}^{n-1} \int_{E_j} E[f(\lambda_j;\chi_t)]dt \qquad (13)$$

since $f(x;\chi_t) \to \chi_t - \hat{\chi}_t$ as $x \to \infty$ so that

$$\int_{E_n} E[f(\lambda_n;\chi_t)]dt = \int_{E_n} E[\chi_t - \hat{\chi}_t]dt = 0 .$$

$I^T[\theta,Y]$ is given by (13) for any choice of θ and χ so

$$\mathbb{C} = \frac{1}{T}\sum_{j=1}^{n-1}\sum_{i=1}^{l} | E_j \cap F_i | \, C(\lambda_j,c_i) .$$

$C(\lambda_n,c_i) = C(\infty,c_i) = 0$ so

$$\mathbb{C} = \frac{1}{T}\sum_{j=1}^{n}\sum_{i=1}^{l} | E_j \cap F_j | \, C(\lambda_j,c_i) = \frac{1}{T}\int_0^T C(\tilde{\lambda}(t),\tilde{c}(t))dt .$$

Lemma 3: Let the channel noise intensity $\lambda(t)$ be a nonnegative extended real-valued Lebesgue measurable function and define L to be the set of all simple nonnegative extended real-valued Lebesgue measurable functions. Write $\mathbb{C} = \mathbb{C}(\lambda)$ to show the dependence of the channel capacity on $\lambda(t)$. Then

$$\sup_{\substack{\tilde{\lambda} \geq \lambda \\ \tilde{\lambda} \in L}} \mathcal{C}(\tilde{\lambda}) \leq \mathcal{C}(\lambda) \leq \inf_{\substack{\tilde{\lambda} \leq \lambda \\ \tilde{\lambda} \in L}} \mathcal{C}(\tilde{\lambda}) .$$

Proof: We prove only the first inequality; the second is shown similarly. Write $I^T[\theta,Y] = I^T[\theta,Y;\lambda]$ to show the dependence of the average mutual information on $\lambda(t)$. According to (3),

$$I^T[\theta,Y;\lambda] = \int_0^T E[f(\lambda(t);\chi_t)]dt .$$

$f(x;\chi_t)$ is a nonincreasing function of x since

$$\frac{\partial f(x;\chi_t)}{\partial x} = \ln\left(\frac{x+\chi_t}{x+\hat{\chi}_t}\right) - \left(\frac{x+\chi_t}{x+\hat{\chi}_t} - 1\right) \leq 0 \quad \text{for all } x \geq 0 .$$

Consequently $I^T[\theta,Y;\lambda] \geq I^T[\theta,Y,\tilde{\lambda}]$ for any $\tilde{\lambda} \in L$ with $\tilde{\lambda} \geq \lambda$ on $[0,T]$. Therefore

$$\frac{1}{T} I^T[\theta,Y;\lambda] \geq \sup_{\substack{\tilde{\lambda} \geq \lambda \\ \tilde{\lambda} \in L}} \frac{1}{T} I^T[\theta,Y;\tilde{\lambda}]$$

and

$$\mathcal{C}(\lambda) = \sup_{\theta,\chi} \frac{1}{T} I^T[\theta,Y;\lambda] \geq \sup_{\theta,\chi} \sup_{\substack{\tilde{\lambda} \geq \lambda \\ \tilde{\lambda} \in L}} \frac{1}{T} I^T[\theta,Y;\tilde{\lambda}]$$

$$= \sup_{\substack{\tilde{\lambda} \geq \lambda \\ \tilde{\lambda} \in L}} \sup_{\theta,\chi} \frac{1}{T} I^T[\theta,Y;\tilde{\lambda}]$$

$$= \sup_{\substack{\tilde{\lambda} \geq \lambda \\ \tilde{\lambda} \in L}} \mathcal{C}(\tilde{\lambda}) .$$

Theorem 4: Let the output $Y = \{Y_t\}_{0 \leq t \leq T}$ of a Poisson channel be directed by $\eta_t = \lambda(t) + \chi_t$ where the channel noise intensity $\lambda(t)$ is deterministic, nonnegative, and Lebesgue measurable and where $\chi_t = \chi_t(\theta,Y)$ is $\mathbf{F}_t^\theta \vee \mathbf{F}_t^Y$ -adapted. Suppose, for some Lebesgue-measurable function $c(t)$, $0 \leq \chi_t \leq c(t) \leq M$ for all $t \in [0,T]$ and some $M < \infty$. Then

$$\mathcal{C} = \frac{1}{T} \int_0^T C(\lambda(t), c(t)) dt .$$

Proof: Write $\mathcal{C} = \mathcal{C}(\lambda, c)$ to make explicit the dependence of $\mathcal{C}$ on $\lambda(t)$ and $c(t)$. By Lemmas 2 and 3,

$$\mathcal{C}(\lambda,c) \le \inf_{\substack{\tilde\lambda \le \lambda \\ \tilde\lambda \in L}} \mathcal{C}(\tilde\lambda, c)$$

$$\le \inf_{\substack{\tilde\lambda \le \lambda \\ \tilde\lambda \in L}} \inf_{\substack{\tilde c \ge c \\ \tilde c \in L}} \mathcal{C}(\tilde\lambda, \tilde c)$$

$$= \inf_{\substack{\tilde\lambda \le \lambda \\ \tilde\lambda \in L}} \inf_{\substack{\tilde c \ge c \\ \tilde c \in L}} \frac{1}{T}\int_0^T C(\tilde\lambda(t), \tilde c(t))dt .$$

C is bounded so $C(\tilde\lambda(t), \tilde c(t))$ is a bounded simple function if $\tilde\lambda \in L$ and $\tilde c \in L$. Also $C(x, y)$ is strictly decreasing in x and strictly increasing in y so $\tilde\lambda \le \lambda$, $\tilde c \ge c$ implies $C(\tilde\lambda, \tilde c) \ge C(\lambda, c)$. Define $h(t) = C(\lambda(t), c(t))$ and $\tilde h = C(\tilde\lambda(t), \tilde c(t))$. Then

$$\mathcal{C}(\lambda,c) \le \inf_{\substack{\tilde h \ge h \\ h \in L}} \frac{1}{T}\int_0^T \tilde h(t)dt .$$

$h = C(\lambda, c)$ is Lebesgue-measurable and bounded so

$$\inf_{\substack{\tilde h \ge h \\ \tilde h \in L}} \frac{1}{T}\int_0^T \tilde h(t)dt = \frac{1}{T}\int_0^T h(t)dt$$

(p. 79 of [14]). Thus

$$\mathcal{C} \le \frac{1}{T}\int_0^T C(\lambda(t))dt .$$

Likewise, starting with the first inequality in Lemma 3 it can be shown that

$$\mathcal{C} \ge \frac{1}{T}\int_0^T C(\lambda(t))dt$$

and the result follows.

Theorem 5: Let the noise and jamming intensities of a Poisson channel be deterministic, nonnegative, and Lebesgue measurable: $\Lambda_t = \lambda(t)$ and $J_t = J(t)$ as in (9). Let the channel encoder output be peak-constrained and let the jamming intensity be average-constrained,

$$\frac{1}{T}\int_0^T J(t)dt \leq P , \tag{14}$$

$P > 0$. Then the optimal choice of jamming intensity is $J_{opt}(t) = J_o(t)$ a.e. for

$$J_o(t) = [\alpha - \lambda(t)]^+$$

(with α defined as in (10).)

Proof: Write $C = C(\lambda+J)$ to show the dependence of the channel capacity on $\lambda(t) + J(t)$. We shall show that

$$C(\lambda+J_o) \leq C(\lambda+J) \tag{15}$$

for all deterministic, nonnegative, Lebesgue measurable J satisfying (14).

In showing (15), only J satisfying

$$\frac{1}{T}\int_0^T J(t)dt = P \tag{16}$$

need be considered. This is so since if

$$\frac{1}{T}\int_0^T J(t)dt < P$$

then there exists a jamming intensity $J_1(t)$ with

$$\frac{1}{T}\int_0^T J_1(t)dt = P$$

such that $J_1 \geq J$ on $[0,T]$ giving $C(\lambda+J_1) \leq C(\lambda+J)$. So we take (16) to hold.

Let f_1, f_2 be nonnegative extended real-valued Lebesgue measurable functions on $[0,T]$ such that $f_1 \leq f_2$ and f_1, f_2 are real-valued on $\{t : f_1(t) < f_2(t)\}$. Define $A \equiv \{(t,z): f_1(t) < z \leq f_2(t)\}$. Define the set function ξ such that $\xi(A) = C(f_1) - C(f_2)$. ξ

extends to a measure on (X,S) where $X = [0,T]\times[0,\infty)$ and S is the σ-field of two-dimensional Lebesgue-measurable subsets of X. ξ hereafter denotes this extension.

Define $x(t) = [\lambda(t)+J_o(t)] \wedge [\lambda(t)+J(t)]$. Also define

$$A_a = \{(t,z): x(t) < z \le \lambda(t)+J(t)\},$$

$$A_b = \{(t,z): x(t) < z \le \lambda(t)+J_o(t)\}.$$

Let m be two-dimensional Lebesgue measure on (X,S) and let $\overline{m}$, $\overline{\xi}$ be the restrictions of m, ξ, respectively, to $A_a \cup A_b$. $\overline{m}$ and $\overline{\xi}$ are then equivalent finite measures. Also

$$\mathcal{C}(\lambda+J) = \mathcal{C}(x) - \overline{\xi}(A_a),$$

$$\mathcal{C}(\lambda+J_o) = \mathcal{C}(x) - \overline{\xi}(A_b).$$

To complete the proof we show $\overline{\xi}(A_a) \le \overline{\xi}(A_b)$.

By the Approximation Theorem, there are finite sequences

$$\{R_1^a, R_2^a, \cdots, R_{n_a}^a\},$$

$$\{R_1^b, R_2^b, \cdots, R_{n_b}^b\}$$

of disjoint closed rectangles, all identically sized,

$$R_i^a = [t_i^a, t_i^a + \Delta_t] \times [z_i^a, z_i^a + \Delta_z],$$

$$R_j^b = [t_j^b, t_j^b + \Delta_t] \times [z_j^b, z_j^b + \Delta_z], \qquad (17)$$

$\Delta_t > 0$, $\Delta_z > 0$, with $R_i^a \subset A_a$, $R_j^b \subset A_b$ for all i, j such that

$$\overline{\xi}(A_a) \le \sum_{i=1}^{n_a} \overline{\xi}(R_i^a) + \delta,$$

$$\overline{\xi}(A_b) \le \sum_{j=1}^{n_b} \overline{\xi}(R_j^b) + \delta$$

for $\delta > 0$ as small as desired. $\overline{m}$ is finite and $\overline{m} \ll \overline{\xi}$ so, for δ small enough,

$$\overline{m}(A_a) \le \sum_{i=1}^{n_a} \overline{m}(R_i^a) + \epsilon,$$

$$\overline{m}(A_b) \le \sum_{j=1}^{n_b} \overline{m}(R_j^b) + \epsilon$$

for any $\epsilon > 0$. Let $n = n_a \wedge n_b$. Then

$$\overline{\xi}(A_a) - \delta_o - \sum_{i=n+1}^{n_a} \overline{\xi}(R_i^a) = \sum_{i=1}^{n} \overline{\xi}(R_i^a) , \tag{18}$$

$$\overline{\xi}(A_b) - \delta_{oo} - \sum_{j=n+1}^{n_b} \overline{\xi}(R_j^b) = \sum_{j=1}^{n} \overline{\xi}(R_j^b) \tag{19}$$

where $0 \le \delta_o \le \delta$, $0 \le \delta_{oo} \le \delta$, and

$$\overline{m}(A_a) - \epsilon_o - \sum_{i=n+1}^{n_a} \overline{m}(R_i^a) = \sum_{i=1}^{n} \overline{m}(R_i^a) , \tag{20}$$

$$\overline{m}(A_b) - \epsilon_{oo} - \sum_{j=n+1}^{n_b} \overline{m}(R_j^b) = \sum_{j=1}^{n} \overline{m}(R_j^b) \tag{21}$$

where $0 \le \epsilon_o \le \epsilon$, $0 \le \epsilon_{oo} \le \epsilon$.

With the rectangles R_i^a, R_j^b, $1 \le i \le n_a$, $1 \le j \le n_b$, given as in (17), we can write

$$\overline{\xi}(R_i^a) = \Delta_t \left[C(z_i^a) - C(z_i^a + \Delta_z , c) \right] ,$$

$$\overline{\xi}(R_j^b) = \Delta_t \left[C(z_j^b) - C(z_j^b + \Delta_z , c) \right] .$$

Also, by the form of $J_o(t)$,

$$\text{essinf}\{ z : (t, z) \in A_a \} \ge \text{esssup}\{ z : (t, z) \in A_b \}$$

so that $z_{j_b} + \Delta_z \le z_i^a$ for all i, j. The partial derivative of $C(x, c)$ w.r.t. x is strictly increasing so

$$C(z_j^b, c) - C(z_j^b + \Delta_z , c) > C(z_i^a, c) - C(z_i^a + \Delta_z , c) .$$

Thus, for all i, j,

$$\overline{\xi}(R_j^b) > \overline{\xi}(R_i^a) .$$

Thus (18) and (19) can be combined:

$$\overline{\xi}(A_b) - \overline{\xi}(A_a) \ge \delta_{oo} - \delta_o + \sum_{j=n+1}^{n_b} \overline{\xi}(R_j^b) - \sum_{i=n+1}^{n_a} \overline{\xi}(R_i^a) . \tag{22}$$

$\overline{m}(R_i^a) = \overline{m}(R_j^b)$ for all i, j so from (20) and (21) comes

$$\overline{m}(A_b) - \overline{m}(A_a) = \epsilon_{oo} - \epsilon_o + \sum_{j=n+1}^{n_b} \overline{m}(R_j^b) - \sum_{i=n+1}^{n_a} \overline{m}(R_i^a) .$$

Also $\overline{m}(A_a) = \overline{m}(A_b)$ since

$$\int_0^T J(t)dt = \int_0^T J_o(t)dt .$$

So

$$\sum_{j=n+1}^{n_b} \overline{m}(R_j^b) - \sum_{i=n+1}^{n_a} \overline{m}(R_i^a) = \epsilon_o - \epsilon_{oo} .$$

Suppose $n = n_a$. Then

$$\sum_{i=n+1}^{n_a} \overline{m}(R_i^a) = 0 , \qquad \sum_{i=n+1}^{n_a} \overline{\xi}(R_i^a) = 0$$

so that

$$\sum_{j=n+1}^{n_b} \overline{m}(R_j^b) = \epsilon_o - \epsilon_{oo} \leq \epsilon .$$

Then, for any $\gamma > 0$, by making ϵ small enough,

$$\sum_{j=n+1}^{n_b} \overline{\xi}(R_j^b) \leq \gamma$$

since $\overline{\xi}$ is finite and $\overline{\xi} \ll \overline{m}$. Then (22) becomes

$$\overline{\xi}(A_b) - \overline{\xi}(A_a) \geq \delta_{oo} - \delta_o - \gamma .$$

δ_{oo}, δ_o, γ may all be made arbitrarily small so $\overline{\xi}(A_b) \geq \overline{\xi}(A_a)$.

Now suppose $n = n_b$. Then

$$\sum_{j=n+1}^{n_b} \overline{\xi}(R_j^b) = 0$$

and

$$\overline{\xi}(A_b) - \overline{\xi}(A_a) \geq \delta_{oo} - \delta_o + \sum_{i=n+1}^{n_a} \overline{\xi}(R_i^a) \geq \delta_{oo} - \delta_o .$$

Again, $\overline{\xi}(A_b) \geq \overline{\xi}(A_a)$.

Theorem 6: Suppose the encoder intensity χ_t of a Poisson channel is $F_t^\theta \vee F_t^Y$-adapted and peak-constrained $0 \leq \chi_t \leq c(t)$ with $c(t) \in \Gamma$ where

$$\Gamma = \{c(t) \geq 0: \frac{1}{T}\int_0^T c(t)dt \leq P\}.$$

Also suppose the channel noise intensity is nonrandom $\Lambda_t = \lambda(t)$. Then the channel capacity is

$$\mathbb{C} = \frac{P}{e}.$$

Proof: By Theorem 4,

$$\mathbb{C} = \sup_{c \in \Gamma} \frac{1}{T}\int_0^T C(\lambda(t), c(t))dt.$$

Define

$$\Gamma_1 = \{c(t) \geq 0: \frac{1}{T}\int_0^T c(t)dt = P\}.$$

Since $C(x,y)$ is increasing in y,

$$\mathbb{C} = \sup_{c \in \Gamma_1} \frac{1}{T}\int_0^T C(\lambda(t), c(t))dt.$$

$C(x,y)$ is decreasing in x and $C(0,y) = y/e$ so

$$\mathbb{C} \leq \sup_{c \in \Gamma_1} \frac{1}{T}\int_0^T C(0, c(t))dt = \sup_{c \in \Gamma_1} \frac{1}{T}\int_0^T \frac{c(t)}{e}dt = \frac{P}{e}.$$

Next we show $\mathbb{C} \geq P/e$ to complete the proof. Choose $L \geq 0$ and let $G = \{t \in [0,T]: \lambda(t) < L\}$. Define $\lambda_L(t) = L$ on G and $\lambda_L(t) = \infty$ on $[0,T]/G$. Then $C(\lambda(t), c(t)) \geq C(\lambda_L(t), c(t))$ so

$$\mathbb{C} \geq \sup_{c \in \Gamma_1} \sup_{L > 0} \frac{1}{T}\int_0^T C(\lambda_L(t), c(t))dt$$

$$= \sup_{c \in \Gamma_1} \sup_{L > 0} \frac{1}{T} \int_G C(L, c(t))\,dt$$

$$= \sup_{L > 0} \sup_{c \in \Gamma_1} \frac{1}{T} \int_G C(L, c(t))\,dt \ .$$

Let S be the set of all nonnegative (Lebesgue-measurable) simple functions on $[0, T]$. Then

$$C \geq \sup_{L > 0} \sup_{c \in \Gamma_1} \sup_{\substack{\tilde{c} \in S \\ \tilde{c} \leq c}} \frac{1}{T} \int_G C(L, \tilde{c}(t))\,dt$$

$$= \sup_{L > 0} \sup_{\tilde{c} \in \Gamma_1 \cap S} \frac{1}{T} \int_G C(L, \tilde{c}(t))\,dt$$

$$= \sup_{L > 0} \sup_{M \geq P} \sup_{\substack{\tilde{c} \in \Gamma_1 \cap S \\ \tilde{c} \leq M}} \frac{1}{T} \int_G C(L, \tilde{c}(t))\,dt \ .$$

$C(x, y)$ as a function of y is strictly convex and strictly increasing for $x > 0$ so

$$\sup_{\substack{\tilde{c} \in \Gamma_1 \cap S \\ \tilde{c} \leq M}} \frac{1}{T} \int_G C(L, \tilde{c}(t))\,dt = \frac{1}{T} \int_G C(L, \tilde{c}_M(t))\,dt$$

where $\tilde{c}_M(t) = M 1_A(t)$ with $|A| = \frac{P}{M} |G|$. Then

$$C \geq \sup_{L > 0} \sup_{M \geq P} \frac{1}{T} \int_G C(L, \tilde{c}_M(t))\,dt$$

$$= \sup_{L > 0} \sup_{M \geq P} \frac{1}{T} C(L, M) \frac{P}{M} |G|$$

$$= P \sup_{L > 0} \frac{|G|}{T} \sup_{M \geq P} \frac{1}{M} C(L, M)$$

$$= P \sup_{L > 0} \frac{|G|}{T} \sup_{M \geq P} C(L/M, 1)$$

$$= P \sup_{L > 0} \frac{|G|}{T} \lim_{M \to \infty} C(L/M, 1)$$

$$= P \sup_{L > 0} \frac{|G|}{T} C(0,1) \ .$$

$C(0,1) = 1/e$ and $|G| \to T$ as $L \to \infty$ so the proof is complete.

AKNOWLEDGEMENT

The author wishes to thank Prof. C.R. Baker for proposing the problem and for suggesting many improvements to the manuscript.

VI. References

[1] P. Brémaud, *Point Processes and Queues*, Springer-Verlag, New York, 1981.

[2] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer-Verlag, New York, 1988.

[3] M. Davis, "Capacity and cut-off rate for Poisson-type channels," *IEEE Transactions on Information Theory*, Vol. 26, No. 6, pp. 710-715, 1980.

[4] S. Forrest, "Optical detectors: three contenders," *IEEE Spectrum*, Vol. 23, No. 5, May 1986.

[5] P. Hoel, S. Port, and C. Stone, *Introduction to Stochastic Processes*, Houghton Mifflin, Co., Boston, 1972.

[6] E. Hoversten, D. Snyder, R. Harger, and K. Kurimoto, "Direct-Detection Optical Communication Receivers," *IEEE Trans. Commun.*, Vol. 22, No. 1, 1974.

[7] Y. Kabanov, "The capacity of a channel of the Poisson type," *Theory of Probability and Applications*, Vol. 23, 1974.

[8] S. Karp and J. Clark, "Photon Counting: A Problem in Classical Noise Theory," *IEEE Trans. Info. Theory*, Vol. 16, No. 6, Nov. 1970.

[9] R. Liptser and A. Shiryayev, *Statistics of Random Processes, Vol. II, Applications*, Springer-Verlag, New York, 1977.

[10] J. O'Reilly and J. da Rocha, "Improved Error Probability Evaluation Methods for Direct Detection Optical Communication Systems," *IEEE Trans. Info. Theory*, Vol. 33, No. 6, November 1987.

[11] S. Personick, "Receiver Design for Digital Optic Communications Systems, I," *Bell Syst. Tech. Jour.*, Vol. 52, No. 6, 1973.

[12] S. Personick, "Receiver Design for Digital Optic Communications Systems, II," *Bell Syst. Tech. Jour.*, Vol. 52, No. 6, 1973.

[13] B. Reiffen and H. Sherman, "An Optimum Demodulator for Poisson Processes: Photon Source Detectors," *Proc. IEEE*, Vol. 51, pp. 1366-1322, October 1963.

[14] H. Royden, *Real Analysis*, Macmillan, 1968.

[15] D.L. Snyder, *Random Point Processes*. John Wiley & Sons, Inc., 1975.

[16] S. Srinivasan, "Stochastic point processes and statistical physics," *Jour. Math. Phy. Sci.*, Vol. 5, 1971.