Capacity and Decoding Rules for the Poisson Arbitrarily Varying Channel.

Shraga I. Bross
Department of Electrical Engineering
Technion, Haifa 32000
ISRAEL

Shlomo Shamai (Shitz)
Department of Electrical Engineering
Technion, Haifa 32000
ISRAEL

September 13, 2000

Abstract

The single-user and two-user (multiple-access) Poisson arbitrarily varying channel with input and state (peak and average-power) constraints, but unlimited in bandwidth, is considered. For both cases the deterministic and random code capacity with the average probability of error criterion is obtained. In the single-user case Wyner's [19] decoder attains the deterministic-code capacity whereas for the two-user case a "minimum distance" decoder that belongs to the class of β-decoders [4] is shown to attain the deterministic-code capacity region as claimed.

Index Terms – Arbitrarily varying channel, Poisson channels, capacity, optical CDMA.

I. Introduction

The discrete memoryless arbitrarily varying channel (AVC) models a communication scenario in which a channel parameter may vary with time without memory in an arbitrary and unknown manner during the transmission of a codeword. In this paper it is assumed that the sequence of channel states is selected arbitrarily subject to a constraint - the state average-power constraint, and possibly depending on the codebook (as the state generator is assumed to be cognizant of the code) but independently of the codeword actually sent.

AVC's capacity depends on whether or not random codes are permitted, and whether the average or the maximum probability of error criterion is used. The random coding capacity admits a simple characterization as a min-max of mutual information, unfortunately random codes are not always implementable.

Unless stated otherwise, the term capacity will hereafter always refer to the capacity for *deterministic codes* and the *average probability of error criterion* (random-code capacity is the same under both criteria only in the single-user case). In the absence of state constraints, Ahlswede [1] proved the following *dichotomy :* the capacity equals either zero or else equals the random code capacity. The necessary and sufficient conditions for positive capacity, as well as capacity under a state constraint for the single-user AVC, have been determined by Csiszár-Narayan in [3], wherein it was shown that Ahlswede's alternatives do not necessarily hold under a state constraint. The two-user multiple-access AVC with and without state constraints has been considered by Gubner in [9, 10], wherein necessary conditions for non-empty interior of the capacity region are given. In [2] Ahlswede-Cai settle Gubner's conjecture and prove that the necessary conditions presented in [9, 10] are also sufficient for the multiple-access AVC to have a non-empty interior of the capacity region. For a comprehensive treatment of the discrete AVC the reader is referred to [7, ch. 6] and the Lapidoth-Narayan tutorial [14] where the necessary terminology and major results are summarized.

In this paper, we determine the deterministic-code capacity of the Poisson single-user, and two-user multiple-access, AVC formally defined as follows. The two-user Poisson multiple-access AVC (MAVC) is modeled by two independent users that generate inputs $\lambda_{m_i}(t)$, $i = 1, 2$ and an independent (arbitrarily varying) state input $\lambda_{s_1}(t)$, $0 \le t \le \infty$, that determine the rates of the corresponding doubly stochastic Poisson processes $d_i(t)$ and $s_1(t)$. Specifically, $d_i(t)$ (resp. $s_1(t)$) corresponds to the number of counts registered by a direct detection device (usually a p-i-n diode) in the interval $[0, t]$, in reaction to the input $\lambda_{m_i}(t)$ (resp. $\lambda_{s_1}(t)$) which is proportional to the squared magnitude of the optical field impinging on the detector at time t integrated over its active surface. The observation is

$$\nu(t) = \sum_{i=1}^{2} d_i(t) + s_1(t) + D(t) \ , \tag{1}$$

which is also a Poisson process with instantaneous rate $\lambda_0 + \lambda_{s_1}(t) + \sum_{i=1}^{2} \lambda_{m_i}(t)$. The dark current represented by $D(t)$ is a homogeneous Poisson process of rate λ_0. We adopt an input constraint $\Gamma = (\sigma_1, \sigma_2)$ and state constraint Λ, namely the permissible average power of user i , $i = 1, 2$ is at most $\sigma_i A$ and that of the state input is at most ΛA, where A denotes the users' transmitter peak power, but we do not impose any bandwidth constraints on the users' and state inputs.

The Poisson single-user AVC corresponds to the case dealing with just one user that generates an input $\lambda_m(t)$ and in result the observation is

$$\nu(t) = d(t) + s_1(t) + D(t) \tag{2}$$

The single-user non-AVC capacity, with constant peak constraint imposed on the encoder output, has been obtained by Kabanov [16] and extended by Davis [17] to reflect the imposition of an average-power constraint on the encoder output. In his seminal contribution [19, 20] Wyner found the exact form of the reliability function for this single-user case. The channel capacity for the two-user non-AVC model was found by Lapidoth-Shamai [15] using the Kabanov/Davis formula and therein it is shown that CDMA is optimal in this case, while the exact error exponents for this case have been determined in [21].

The remainder of this paper is organized as follows. In section II we introduce the terminology and summarize our main results. These results are proved in sections III and IV.

II. Definitions and Results

Let $\nu_0^T = \{\nu(t) : 0 \leq t \leq T\}$ be the observation process. A two-user code with parameters $(M_1, M_2, T, \Gamma, P_e)$ is defined by:

- a set of M_1 waveforms (the first user codebook) $\lambda_{m_1}(t), 0 \leq t \leq T, 1 \leq m_1 \leq M_1$ which satisfy the peak and average power constraints $0 \leq \lambda_{m_1}(t) \leq A$, $1/T \int_0^T \lambda_{m_1}(t)dt \leq \sigma_1 A$.
- a set of M_2 waveforms (the second user codebook) $\lambda_{m_2}(t), 0 \leq t \leq T, M_1+1 \leq m_2 \leq M_1 + M_2$ which satisfy the peak and average power constraints $0 \leq \lambda_{m_2}(t) \leq A$, $1/T \int_0^T \lambda_{m_2}(t)dt \leq \sigma_2 A$.
- a "decoder" mapping $D : \{\nu_0^T\} \to (\{1, 2, \ldots, M_1\}, \{M_1+1, \ldots, M_1+M_2\})$.

The average probability of error for this code, when the state input is $\lambda_{s_1}(t)$, is

$$P_e(s_1) = \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=M_1+1}^{M_1+M_2} \Pr\Big\{D(\nu_0^T) \neq (m_1, m_2) | \lambda_{m_1}(\cdot), \lambda_{m_2}(\cdot), \lambda_{s_1}(\cdot)\Big\} ,$$

and the restriction of this definition to the single-user case is obvious.

Assuming that the same peak constraint A holds for both the signal and state seems to be somewhat restrictive hence (in the single-user case) we assume that the state is peak-limited to B and it is allowed to vary arbitrarily within a codeword interval $[0, T]$ subject to the constraints

$$0 \leq \lambda_{s_1}(t) \leq B \ , \ \frac{1}{T} \int_0^T \lambda_{s_1}(t)dt \leq \Lambda A \ . \tag{3}$$

The following result due to Frey summarizes the optimal jamming strategy in the case at hand.

Proposition 1 [18, Theorems 4,5] : For the Poisson channel defined by (2) with non-random noise intensity (dark current) $\lambda_0(t) = \lambda_0$,

- The information capacity (in the sense of the average mutual information between the channel input and output processes over the interval $[0, T]$) is minimized by a *deterministic* jamming intensity $\lambda_{s_1}(t)$.
- If the channel encoder is peak-constrained $0 \leq \lambda_m(t) \leq A$ and the jamming intensity is average-power constrained as in (3), the jamming intensity which minimizes the information capacity is given by the "waterpouring" strategy

$$\lambda_{s_1}^{opt}(t) = [\alpha A - \lambda_0]_+ \ ,$$

where $[z]_+ = \max\{0, z\}$ and α is defined by

$$\int_0^T [\alpha A - \lambda_0]_+ \, dt = \Lambda A T \ .$$

The optimal waveform is thus $\lambda_{s_1}^{opt}(t) = \Lambda A$, i.e. the optimal jamming sequence admits the form of a dark current process, mutually independent of λ_0, with intensity ΛA.

Our main result for the single-user case is the following.

Theorem 1: *The deterministic-code capacity of the Poisson AVC with input constraint $\Gamma = \sigma$, state constraint Λ, and $B \geq A$ is*

$$C = \begin{cases} A f_1(p^*, \Lambda) \ , & \text{if } \ \Lambda \leq p^*(\Lambda, s_0) < \sigma \\ A f_1(\Lambda, \Lambda) \ , & \text{if } \ p^*(\Lambda, s_0) < \Lambda < \sigma \\ 0 \ , & \text{if } \ \Lambda \geq \sigma \end{cases} \tag{4}$$

where $s_0 = \lambda_0 / A$, and $f_1(p,q)$, $p^*(\Lambda, s_0)$ are defined by

$$f_1(p,q) \ \stackrel{\triangle}{=} \ -(s_0 + q + p)\ln(s_0 + q + p) + (1-p)(s_0+q)\ln(s_0+q) + p(s_0+q+1)\ln(s_0+q+1)$$

$$p^*(\Lambda, s_0) \ \stackrel{\triangle}{=} \ \frac{(s_0 + \Lambda + 1)^{s_0 + \Lambda + 1}}{e(s_0 + \Lambda)^{s_0 + \Lambda}} - (s_0 + \Lambda) \ .$$

The random-code capacity is $C_r(\Lambda) = A f_1(p^*, \Lambda)$.
If $B < A$ then $C = C_r(\Lambda)$ (note that $B < A$ implies $\Lambda \leq B/A$).

Average-power constraints in the two-user case (even for the non AVC case) cannot be handled by simply considering the pairs of messages that constantly satisfy the average-power constraint [23, 24, 15], moreover the classic time-sharing principle does not hold for the MAVC subject to a state constraint [9, 10]. Instead, the *random-code* capacity region is expressed in terms of an auxiliary random variable v that can be regarded as a form of time-sharing, and in fact this formulation extends the set of attainable rate pairs. To this end let the channel inputs pdf admit the form

$$P_{V\lambda_1\lambda_2 S_1}(v, \lambda_{m_1}, \lambda_{m_2}, s_1) = \gamma(v)\delta_1(\lambda_{m_1}|v)\delta_2(\lambda_{m_2}|v)r(s_1|v) \ . \tag{5}$$

That is, given a r.v. v (that is indexed by codeword intervals) the users' and state inputs are conditionally independent given $V = v$ and the input/state constraints Γ and Λ are satisfied by input/state waveforms that may sometimes violate the corresponding constraints but when averaged over the auxiliary r.v. they are satisfied;

$$\begin{aligned} E[\ell(S)] \ &- \ \sum_v \gamma(v)\Big(\sum_{s_1} \ell(s_1) r(s_1|v)\Big) = \Lambda \\ E[\ell(\lambda_i)] \ &- \ \sum_v \gamma(v)\Big(\sum_{\lambda_{m_i}} \ell(\lambda_{m_i})\delta_i(\lambda_{m_i}|v)\Big) = \sigma_i \ , \ i = 1, 2 \ . \end{aligned} \tag{6}$$

By the Gubner-Hughes result [11], the random-code capacity region equals the closure of the set of attainable rates via the above strategy.

Herein as we aim at the deterministic-code capacity we consider the set of attainable rates during a portion of time over which the users' and state inputs constantly satisfy a pair of constraints Γ , Λ and furthermore we assume that the users' and state inputs are both peak-limited to A. Let $f(p_1,p_2,q)$ and $g(p_1,p_2,q)$ be defined by

$$\begin{aligned} f(p_1,p_2,q) &\stackrel{\triangle}{=} -(s_0+q+p_1+p_2)\ln(s_0+q+p_1+p_2)+p_1p_2(s_0+q+2)\ln(s_0+q+2) \\ &\quad +(p_1+p_2-2p_1p_2)(s_0+q+1)\ln(s_0+q+1)+(1-p_1)(1-p_2)(s_0+q)\ln(s_0+q) \\ g(p_1,p_2,q) &\stackrel{\triangle}{=} -(1-p_1)(s_0+q+p_2)\ln(s_0+q+p_2)-p_1(s_0+q+1+p_2)\ln(s_0+q+1+p_2) \\ &\quad +p_1p_2(s_0+q+2)\ln(s_0+q+2)+(p_1+p_2-2p_1p_2)(s_0+q+1)\ln(s_0+q+1) \\ &\quad +(1-p_1)(1-p_2)(s_0+q)\ln(s_0+q) \ , \end{aligned}$$

then our result for the two-user case is the following.

Theorem 2: *The deterministic-code capacity region of the Poisson MAVC with constant input constraint $\Gamma=(\sigma_1,\sigma_2)$ and state constraint Λ is*

$$\begin{aligned} C(p_1,p_2,\Lambda) \;=\; \Big\{(R_1,R_2) \;:\; & 0\le R_1\le Ag(p_2,p_1,\Lambda) \ , \\ & 0\le R_2\le Ag(p_1,p_2,\Lambda) \\ & 0\le R_1+R_2\le Af(p_1,p_2,\Lambda)\Big\} \ , \ p_i\le\sigma_i \end{aligned} \tag{7}$$

provided that $\Lambda<\min\{p_1,p_2\}$.
If $\Lambda>\min\{p_1,p_2\}$ the capacity region has an empty interior.
Let

$$\mathcal{R}(\Lambda,\Gamma) \stackrel{\triangle}{=} \bigcup C(p_1,p_2,\Lambda) \ ,$$

where the union is over all sets $|V|\le\infty$ and probability distributions of the form (5) that satisfy (6), then the random-code capacity under constraints, $C_r(\Lambda,\Gamma)$ is equal to the closure of $\mathcal{R}(\Lambda,\Gamma)$ [11].

Actually, we show that the capacities as claimed in Theorems 1 and 2 are achievable using "minimum-distance" decoding rules and in particular the two-user decoding rule belongs to the class of β-decoders defined in [3, 4].

III. Proof of the Single-User Result

Unless otherwise stated we assume that $B=A$.
First we consider a particular case of a single-user code. We fix some small $\Delta>0$ and partition the time interval $[0,T]$ on equal segments of length Δ/A as in [19]. We limit ourselves to input waveforms that take only extreme values 0 or A on each of those segments and the best of such single-user codes gives us the lower bound on the single-user capacity.

Formally, assume the following,

a) The channel input waveforms $\lambda_{m_1}(t),\lambda_{s_1}(t)$ are constant for $(n-1)\Delta/A<t\le n\Delta/A, n=1,2,\ldots$, and take only the values 0 or A. For $n=1,2,\ldots$, let $x_n=0$ or 1 (resp. $x_{s_n}=0$ or 1) according as $\lambda_{m_1}(t)=0$ or A (resp. $\lambda_{s_1}(t)=0$ or A) in the interval $((n-1)\Delta/A,n\Delta/A]$. Consequently, the

receiver observes only samples $\nu(n\Delta/A), n = 1, 2, \ldots$, or alternatively the increments $\hat{y}_n = \nu(n\Delta/A) - \nu((n-1)\Delta/A)$. In the limit as $\Delta \to 0$ - i.e. $\lambda_{m_1}(t)$ (resp. $\lambda_{s_1}(t)$) being an infinitely fast varying process, this assumption does not imply any loss in attainable capacity as justified by the Kabanov-Davis technique [16, 17].

b) The receiver interpretes $\hat{y}_n \geq 2$ as being the same as $\hat{y}_n = 0$. This assumption doesn't harm optimality in our case since $\Pr\{\hat{y}_n \geq 2\} = O(\Delta^2)$ as $\Delta \to 0$ (i.e. $\hat{y}_n \geq 2$ is a rare event when Δ gets small). Thus the receiver has available

$$y_n \triangleq \begin{cases} 1, & \hat{y}_n = 1 \\ 0, & \text{otherwise} \end{cases}$$

Subject to the assumptions a, b the channel reduces to a single-user binary-input binary-output discrete memoreless arbitrarily varying channel with transition probability $W(y|x, s_1)$ given by ($\lambda_0 = s_0 A$)

$$\begin{aligned} W(1|0,0) &- 1 - W(0|00) = s_0\Delta \exp^{-s_0\Delta} = s_0\Delta + O(\Delta^2) \ , \\ W(1|1,0) &- W(1|0,1) = 1 - W(0|1,0) = 1 - W(0|0,1) = \\ &\quad (1+s_0)\Delta \exp^{-(1+s_0)\Delta} = (1+s_0)\Delta + O(\Delta^2) \\ W(1|1,1) &- 1 - W(0|1,1) = (2+s_0)\Delta \exp^{-(2+s_0)\Delta} = (2+s_0)\Delta + O(\Delta^2) \ . \end{aligned} \quad (8)$$

Thus in our channel model the output depends statistically on whether the transmitter, the state, and the dark current source emit zero or at least one emits a one.

A. *Capacity Evaluation via the DMC Decomposition*

Based on our assumption that the channel model (2) can be further decomposed into the discrete-time model (8), in which the input and state alphabets are finite (in fact $|\mathcal{X}| = |\mathcal{S}| = 2$), to compute the capacity of our channel we follow the results of Csiszár [6] and Csiszár-Narayan in [3, 4].

For a state distribution $Q = (1-q, q)$, consider the channel $W_Q : \mathcal{X} \to \mathcal{Y}$

$$\begin{aligned} W_Q(1|0) &= (1-q)s_0\Delta + q(1+s_0)\Delta + O(\Delta^2) = \Delta(s_0 + q) + O(\Delta^2) \\ W_Q(1|1) &= (1-q)(1+s_0)\Delta + q(2+s_0)\Delta + O(\Delta^2) = \Delta(s_0 + q + 1) + O(\Delta^2) \ . \end{aligned}$$

Hence for an input distribution $P = (1-p, p)$ the mutual information $I(P, W_Q)$ denoted by $I(p, q)$, equals

$$I(p,q) = h(\Delta(s_0 + q + p)) - (1-p)h(\Delta(s_0 + q)) - ph(\Delta(s_0 + q + 1))$$

For the binary entropy function $h(u) = -u \ln u - (1-u)\ln(1-u)$ we have

$$\begin{aligned} h(u) &- -u \ln u + u + O(u^2) \ , \\ h(u + O(u^2)) &= h(u) + O(u^2 \ln u) \ , \\ h(Ku) &= -Ku \ln u - Ku \ln K + Ku + O(u^2) \ , \quad u \to 0 \ . \end{aligned}$$

Since later we assume that $\Delta \to 0$ we shall henceforth be interested in our computations up to $O(\Delta)$. It is straightforward to show that

$$I(p;q) = \Delta f_1(p,q) \quad (9)$$

Furthermore, it can be easily verified by differentiation and using Jensen's inequality, that $I(p,q)$ is a decreasing function of q, then $\ell(Q) = q$ implies

$$I(P,\Lambda) \stackrel{\Delta}{=} \min_{q\leq\Lambda} I(p,q) = I(p,\Lambda) \tag{10}$$

Next we determine $\Lambda_0(P)$, namely the minimum of

$$\sum P(x)U(s_1|x)\ell(s_1) = (1-p)U(1|0) + pU(1|1)$$

over the set of channels $U : \mathcal{X} \to \mathcal{S}$ that render the AVC symmetrizable, i.e

$$\sum_{s_1\in\mathcal{S}} W(y|x,s_1)U(s_1|x') = \sum_{s_1\in\mathcal{S}} W(y|x',s_1)U(s_1|x) \ . \tag{11}$$

It suffices to consider (11) with $y = 1, x = 0, x' = 1$, then

$$\sum_{s_1\in\mathcal{S}} W(1|0,s_1)U(s_1|1) = \sum_{s_1\in\mathcal{S}} W(1|1,s_1)U(s_1|0) \ ,$$

implies the relation

$$s_0U(0|1) + (1+s_0)U(1|1) = (1+s_0)U(0|0) + (2+s_0)U(1|0)$$

which is satisfied by exactly those channels U for which $U(1|1) = 1$ and $U(0|0) = 1$. That is, the state strategy that symmetrizes the channel is the one that chooses state inputs identical to the message inputs. Thus $\Lambda_0(P) = p$, $\Lambda_0 = \max_P \Lambda_0(P) = 1$, and since [4]

$$C(\Lambda) = \max_{P:\Lambda_0(P)\geq\Lambda} I(P,\Lambda) \quad , \quad \text{if} \ \ \Lambda < \Lambda_0$$

we obtain for the deterministic-code capacity

$$C(\Lambda) = \max_{p:p\geq\Lambda} I(p,\Lambda) \tag{12}$$

On the other hand the random-code capacity $C_r(\Lambda)$ equals $\max_p I(p,\Lambda)$ with no constraint on p.

Akin to the interpretation of the capacity expression for the ordinary arbitrarily varying OR channel [4, section III], equation (9) offers a geometric interpretation which can be used to obtain $I(p,\Lambda)$ as shown in Fig. 1. Let A, B and D be points on the $t\log t$ curve with $t_1 = s_0 + q$, $t_2 = s_0 + q + 1$, and $t = s_0{+}q{+}p$ respectively. Then $I(p,\Lambda)$ equals that part of the ordinate of D which lies below the secant AB. Further, with increasing p, the point D moves from left to right. $I(p,\Lambda)$ attains its maximum iff the tangent to the $t\log t$ curve at D is parallel to the secant AB. Denoting the maximizing p by $p^* = p^*(\Lambda, s_0)$ we have

$$\begin{aligned}
\frac{\partial I(p,\Lambda)}{\partial p} &- \ln(s_0+\Lambda+p) - 1 - (s_0+\Lambda)\ln(s_0+\Lambda) + (s_0+\Lambda+1)\ln(s_0+\Lambda+1) \\
p^* &- \frac{(s_0+\Lambda+1)^{s_0+\Lambda+1}}{e(s_0+\Lambda)^{s_0+\Lambda}} - (s_0+\Lambda) \\
C_r(\Lambda, s_0) & \quad I(p^*,\Lambda)
\end{aligned}$$

In summary

$$\begin{array}{lll} C(\Lambda) & C_r(\Lambda) & \text{if } \Lambda \leq p^*(\Lambda, s_0) \\ C(\Lambda) & I(\Lambda,\Lambda) < C_r(\Lambda) & \text{if } \Lambda > p^*(\Lambda, s_0) \ . \end{array} \tag{13}$$

Geometrically, the first or second case is obtained based on whether the tangent to the $t \log t$ curve at the point E, corresponding to $t_3 = s_0 + 2\Lambda$, has a smaller (possibly equal) or a larger slope, respectively than the secant AB. In other words, $\Lambda \leq p^*(\Lambda, s_0)$ iff

$$s_0 + 2\Lambda \leq \frac{(s_0 + \Lambda + 1)^{s_0+\Lambda+1}}{e(s_0 + \Lambda)^{s_0+\Lambda}}$$

This completes the proof of the direct part of Theorem 1 when $\Gamma > \Lambda$ and $B = A$.

B. *Code Construction and Decoding Rule*

To describe a single-user code with parameters $T, M = \lceil e^{RT} \rceil$, in which all code waveforms $\{\lambda_m(t)\}$, $1 \leq m \leq M$ take only the extreme values 0 and A, it suffices to construct the supports $b_m = \{t : \lambda_m(t) = A\}$. Our code is precicely Wyner's code [19] that is - for a given p of the form $p = k/M$, $0 < p < 1$ let $\mathcal{A} = \{a_{m,j}\}$ be the $M \times \binom{M}{k}$ binary matrix, the columns of which are the $\binom{M}{k}$ binary M-vectors with exactly k ones (and $(M-k)$ zeros). Divide the interval $[0, T]$ into $\binom{M}{k}$ subintervals Δ_j, each of length $T/\binom{M}{k}$. Put $\Delta_j \subset b_m$ iff $a_{m,j} = 1$, i.e. the support b_m consists of the union of all subintervals Δ_j such that $a_{m,j} = 1$. As we let $T \to \infty, \lambda_m(t)$ satisfies (henceforth $\mu(B)$ denotes Lebesgue measure of a measurable set B).

$$\begin{array}{rcl} \frac{1}{T}\mu\{t : \lambda_m(t) = A\} & \approx & p \\ \frac{1}{T}\mu\{t : \lambda_{m_1}(t) = A \ , \ \lambda_{m_2}(t) = 0\} & \approx & p(1-p) \ , \ m_1 \neq m_2 \ . \end{array}$$

Decoder : Any hypothesis m defines two independent Poisson processes; the first with transmission rate $\zeta_m(t) = A + \lambda_0 + \lambda_{s_1}(t)$ on the interval b_m where m is "active", and the second with transmission rate $\zeta_m(t) = \lambda_0 + \lambda_{s_1}(t)$ on the interval where only the dark current and state input are active. The conditional sample function density of the observation is given by (e.g. [27, 28, 29])

$$\Pr[\{\nu(t), 0 \leq t \leq T\}|m] = \exp\Big[-\int_0^T \zeta_m(t)dt\Big] \prod_{i=1}^{n} \zeta_m(a_i) \ ,$$

where $\{a_1, \ldots, a_n\}$ is the realization of the observed point process in the interval $[0, T]$.

Since the integral

$$\int_0^T \zeta_m(t)dt = (\lambda_0 + (p+q)A)T \quad , \quad q \leq \Lambda$$

does not depend on the message m, a maximum likelihood message decoder chooses m^* such that [27]

$$m^* \in \arg\ \max \prod_{i=1}^{n} \zeta_m(a_i) \ . \tag{14}$$

Unfortunately, the arbitrarily varying property of the channel prevents the use of the optimum likelihood ratio

$$\frac{dP_{m'}}{dP_m}(a_1,\ldots,a_n) = \prod_{i=1}^{n}\frac{\zeta_{m'}(a_i)}{\zeta_m(a_i)}\ , \tag{15}$$

since the decoder cannot determine the exact intensity at the realization moments. To this end for $1 \le m \le M$ the decoder observes ν_0^T, computes

$$\psi_m \ = \ \int_{b_m} d\nu(t) = \{\text{number of arrivals in } b_m\}\ ,$$

and then $D(\nu_0^T) = m^*$ if

$$\begin{aligned} \psi_m \ &< \ \psi_{m^*} \qquad 1 \le m < m^*\ , \\ \psi_m \ &< \ \psi_{m^*} \qquad m^* \le m < M\ . \end{aligned} \tag{16}$$

That is, ν_0^T is decoded as that m which is "nearest" to the observed realization in the sense that it maximizes ψ_m. By (15) this decoding rule is also a maximum likelihood decoding rule, and its performance has been analyzed in [19] wherin it is shown that as $T \to \infty$

$$\begin{aligned} P_{em} \ &- \ \Pr\Big\{D(\nu_0^T) \neq m|\lambda_m(\cdot) \text{ is transmitted}\Big\} \\ &< \ \exp\Big\{-T\Big[AE_1(\rho,p)|_{s=s_0+q} - \rho R\Big] + o(T)\Big\}\ , \end{aligned}$$

where

$$E_1(\rho,p)|_{s=s_0+q} \ = \ s_0+q+p-\Big[(1-p)(s_0+q)^{\frac{1}{1+\rho}} + p(s_0+q+1)^{\frac{1}{1+\rho}}\Big]^{1+\rho}\ .$$

Since

$$\left.\frac{\partial E_1(\rho,p)|_{s=s_0+q}}{\partial \rho}\right|_{\rho=0} = f_1(p,q)$$

the capacity as claimed in Theorem 1 is achievable by Wyner's encoder/decoder.

The fact that $\Gamma \le \Lambda$ implies $C = 0$ follows by the well-known argument of Blackwell *et. al.* [8]. Namely if $\lambda_1,\ldots,\lambda_N$ is an arbitrary set of N code waveforms satisfying the average-power constraint, then assuming $\Gamma \le \Lambda$ consider the state input sequence that symmetrizes the channel - i.e. $\lambda_{s_1}(t) = \lambda_i, (i-1)T \le t \le iT, 1 \le i \le N$. In this case [5]

$$\begin{aligned} \Pr\{D(\nu_0^T) \neq i|\lambda_i, s_1 = j, \lambda_0\} \ &- \ \Pr\{D(\nu_0^T) \neq i|\lambda_j, s_1 = i, \lambda_0\} \\ &> \ 1 - \Pr\{D(\nu_0^T) \neq j|\lambda_j, s_1 = i, \lambda_0\} \end{aligned}$$

which implies that $1/N\sum_{j=1}^{N} P_e(s_1 = j) \ge 1/4$ and consequently $P_e(s_1 = j) \ge 1/4$ for at least one $j \in \{1,\ldots,N\}$ (see [5, Appendix]).

The cases $B > A$ and $B < A$ are relegated to Appendix I.

C. *Converse Theorem*

Given that the jamming strategy that minimizes the information capacity admits the form of an independent constant-intensity process we provide two alternatives for proving the converse.

Set

$$\mathsf{S}^n(\Lambda) \triangleq \{\lambda_{s_1}(t) \ : \ \lambda_{s_1}(t) = \Lambda' A \ , \ \Lambda' \leq \Lambda \ , \ 0 \leq t \leq T\} \ ,$$

and let C^Λ denote the capacity, under state constraint Λ, as claimed in Theorem 1. It can be verified that if $\Lambda_1 \leq \Lambda_2$ then $C^{\Lambda_2} \leq C^{\Lambda_1}$, and furthermore (cf. [10, section III Lemma 3.1])

$$\bigcap_{0<\delta<\Lambda} C^{\Lambda-\delta} = C^\Lambda \ .$$

Consequently it suffices to show that for every $0 < \delta < \Lambda$, the capacity under state constraint Λ is bounded above by $C^{\Lambda-\delta}$. Suppose $R \in C^\Lambda$ then there exists a code with this rate and a decoder with $P_e \leq \epsilon$ for all $\lambda_{s_1} \in \mathsf{S}^n(\Lambda)$.

Consider a sequence $\boldsymbol{S} = (S_1, \ldots, S_n)$ of samples of a deterministic function $\lambda_{s_1}(t) = (\Lambda - \delta)A$. For any given code with codewords $\lambda_1, \ldots, \lambda_N$ and decoder $D(\nu_0^T)$,

$$\begin{aligned} EP_e(\boldsymbol{S}) \qquad & \frac{1}{N}\sum_{i=1}^{N} \Pr\{D(\nu_0^T) \neq i|\lambda_i, (\Lambda-\delta)A, \lambda_0\} \\ < & \max_{\lambda_{s_1} \in \mathsf{S}^n(\Lambda)} P_e(s_1) + \Pr\left\{\frac{1}{T}\int_0^T \lambda_{s_1}(t)dt > \Lambda A\right\} \ . \end{aligned} \tag{17}$$

Since $\lambda_{s_1}(t) = (\Lambda - \delta)A$ the second term in the r.h.s. of (17) is zero and we obtain

$$\frac{1}{N}\sum_{i=1}^{N} \Pr\{D(\nu_0^T) \neq i|\lambda_i, (\Lambda-\delta)A, \lambda_0\} \leq \epsilon \ .$$

Here the left side is the average error probability of the given code on the "ordinary" Poisson channel with additive dark current of intensity $\lambda_0 + (\Lambda - \delta)A = A(s_0 + \Lambda - \delta)$. Hence, by the converse to the coding theorem for such channels, it follows that $C^{\Lambda-\delta} \leq Af_1(p, \Lambda - \delta)$. ∘

Alternatively, we adopt the approach outlined in [22] and further extended in [21] to obtain a sphere packing lower bound on the error probability by means of which we derive a converse coding theorem. For a given photon arrival count n and any hypothesis of transmitted message m, we consider the corresponding configuration of counts (n_1, n_0) namely the arrival counts on b_m and on b_m^c (b_m^c denotes the complement of b_m on $[0,T]$). Each such configuration is associated with a "volume" which is a function of n_j and $t_j, j = 0, 1$, where $t_1 = p$ and $t_0 = 1 - p$.

Assume that a code $\left\{\lambda_m(\cdot)\right\}$ of cardinality M and a decoding rule $D(\nu_0^T)$ have been chosen. Without loss of generality we may assume that each of the M signals $\lambda_m(\cdot)$ has support of Lebesgue measure p on $[0, T]$.

Remark : The optimality of equi-energy signaling from the error exponent aspect is proved in [20] and [21, Appendix II].

Assume further that the observation space $\mathcal{X}$ is

$$\mathcal{X} = \sum_{n=0}^{\infty} \mathcal{X}_n \ ,$$

where the set $\mathcal{X}_n, n \geq 1$, consists of all sequences $0 \leq x_1 \leq x_2 \ldots \leq x_n \leq T$ and each $x_\ell, \ell = 1, \ldots, n$ represents the photon arriving moment. This implies that to each message $\theta_i, i = 1, \ldots, M$ corresponds a region $\mathcal{D}_{in} \subseteq \mathcal{X}_n, n = 1, \ldots$ of making the decision in its favor, and the decision regions satisfy

$$\bigcup_{i=1} \mathcal{D}_{in} = \mathcal{X}_n \quad \text{and} \quad \mathcal{D}_{in} \cap \mathcal{D}_{jn} = \emptyset \ , \ i \neq j \ .$$

Following [22] introduce the "volume" $|\mathcal{X}_n|$ of the space $\mathcal{X}_n$

$$|\mathcal{X}_n| = \int \cdots \int_{0 \leq x_1 \leq x_2 \ldots \leq x_n \leq T} dx_1 \ldots dx_n = \frac{T^n}{n!} \triangleq v(n, T) \ .$$

If θ_i is transmitted and $\nu(T) = n$ the conditional distribution of the photon counts on the two slots b_m and b_m^c is given by

$$\Pr\left\{\psi_m^1 = n_1, \psi_m^0 = n - n_1 | \theta_i, \mathcal{X}_n\right\} = \binom{n}{n_1} \pi_1^{n_1} (1 - \pi_1)^{n - n_1} \triangleq P(n_1, n) \ , \tag{18}$$

where

$$\pi_1 - \frac{t_1(1 + s_0 + q)}{s_0 + q + p}$$

and

$$\Pr\left\{\nu(T) = n | \theta_i\right\} = (AT)^n \frac{(s_0 + q + p)^n}{n!} e^{-(s_0+q+p)AT} \ .$$

Notice that the conditional density $p(x_1, \ldots, x_n | \theta_i, n_1, n)$ is constant on the corresponding subset of $\mathcal{X}_n$. Thus we associate a corresponding "volume" to each configuration of counts - $(n_1, n_0 = n - n_1)$ - as follows.

$$\begin{aligned} \Big|\mathcal{X}_n \cap [\theta_i, (n_1, n_0)]\Big| &= v(n_1, t_1 T) v(n_0, t_0 T) \\ &- \frac{T^n t_1^{n_1} t_0^{n_0}}{n_1! n_0!} = \binom{n}{n_1} t_1^{n_1} t_0^{n_0} v(n, T) \\ &\triangleq v(t_1, t_0, n_1, n_0, T) \ . \end{aligned} \tag{19}$$

Consider the volume $|\mathcal{D}_{in}|$ of making the decision in favor of θ_i. This volume admits different values based on the corresponding configuration (n_1, n_0) of the given realization of $\mathcal{X}_n$ with respect to the waveform message θ_i. Thus, let $|\mathcal{D}_{in}^{(n_1, n_0)}|$ denote the optimum $|\mathcal{D}_{in}|$ for a realization $\mathcal{X}_n$ having configuration (n_1, n_0) then the conditional error probability (of θ_i) averaged over all possible configurations is lower bounded by

$$P(e|\theta_i, \mathcal{X}_n) \geq \sum_{n_1 + n_0 = n} \left[1 - \frac{|\mathcal{D}_{in}^{(n_1, n_0)}|}{v(t_1, t_0, n_1, n_0, T)}\right]_+ \Pr\{n_1 | \theta_i, \mathcal{X}_n\} \ . \tag{20}$$

This implies that the optimum distribution of $\{|\mathcal{D}_{in}^{(n_1, n_0)}|\}$ that minimizes the r.h.s. of (20) needs to "cover" an average volume that equals at least the conditional expected value of $v(t_1, t_0, n_1, n_0, T)$ w.r.t. the distribution $P(n_1, n)$.

A lower bound on the average error probability is obtained by considering the contribution of all possible messages $\theta_j, j = 1, \ldots, M$ for a given $\mathcal{X}_n$. Thus we should consider for each message the corresponding configuration implied by the realization $\mathcal{X}_n$ and assign accordingly the optimal decision regions $\mathcal{D}_{jn}^{(n_1,n_0)}$. Since all messages are equiprobable and exhibit the same statistical properties we can associate with each message an equivalent volume that equals $E_P v(t_1, t_0, n_1, n_0, T)$ and the optimal decision regions assignment reduces to the issue of covering these expected values of message volumes as economically as possible.

Let the volume $v(n,i)$ associated with the message θ_i be given by

$$\begin{aligned} v(n,i) &- E_P v(t_1, t_0, n_1, n_0, T) \\ &> \left[E_P v^\rho(t_1, t_0, n_1, n_0, T)\right]^{\frac{1}{\rho}} \quad , \quad 0 \leq \rho \leq 1 \end{aligned} \tag{21}$$

where E_P denotes expectation w.r.t. $P(n_1, n)$ and the last step follows by the Jensen inequality.

Following the same procedure as in [21, section IV] one can show that

$$\begin{aligned} \left[E_P v^\rho(t_1, t_0, n_1, n_0, T)^\rho\right]^{\frac{1}{\rho}} &> \left\{\left[\sum_j t_j \left(\frac{\kappa_j}{\sum_j t_j \kappa_j}\right)^{\frac{1}{1+\rho}}\right]^{\frac{1+\rho}{\rho}}\right\}^n v(n,T) \\ &\stackrel{\Delta}{=} \tilde{v}(n,i) v(n,T) \end{aligned} \tag{22}$$

where κ_j denotes the sum of peak intensities of all photon emissions at time slot j - i.e. $\kappa_0 = s_0 + q, \kappa_1 = s_0 + q + 1$, and we note that $\tilde{v}(n,i)$ is in fact independent of i.

Let $|\tilde{\mathcal{D}}_{in}|$ denote the equivalent volume of the decision set $\mathcal{D}_{in}$ associated with the average volume $v(n,i)$ then using this notation the conditional error probability lower bound (20) can be stated as

$$P(e|\theta_i, \mathcal{X}_n) \geq \left[1 - \frac{|\tilde{\mathcal{D}}_{in}|}{\tilde{v}(n,i) v(n,T)}\right]_+ ,$$

It follows that the average error probability $P_e(M, \{t_j^{(i)}\}, \{\kappa_j^{(i)}\}, \{\mathcal{D}_{in}\})$ is lower bounded by

$$P_e(M, \{t_j^{(i)}\}, \{\kappa_j^{(i)}\}, \{\mathcal{D}_{in}\}) \geq \frac{1}{M} \sum_{n=0}^{\infty} \frac{(AT)^n}{n!} \sum_{i=1}^{M} e^{-(\sum_j t_j^{(i)} \kappa_j^{(i)})AT} (\sum_j t_j^{(i)} \kappa_j^{(i)})^n \left[1 - \frac{|\tilde{\mathcal{D}}_{in}|}{\tilde{v}(n,i) v(n,T)}\right]_+ ,$$

where we have associated with message $\theta_i, i = 1, \ldots, M$ the parameters $\{t_j^{(i)}\}, \{\kappa_j^{(i)}\}$.

Define the fractionals of the decision regions d_{in} by $d_{in} = \frac{|\tilde{\mathcal{D}}_{in}|}{v(n,T)}$, then a "good" packing set $\{d_{in}\}$, $n = 0, 1, \ldots$ should satisfy

$$\begin{aligned} d_{in} &< \tilde{v}(n,i) \qquad i = 1, \ldots, M \\ \sum_{i=1}^{M} d_{in} &< 1 \end{aligned}$$

Recall that we've assumed a fixed composition code for each of the users - i.e. $\{t_j^{(i)}\}, \{\kappa_j^{(i)}\} - \{t_j\}, \{\kappa_j\}$ $\forall i$ then the optimal $\{d_{in}\}$ for this case are

$$d_{in} = \min \left\{1/M, \tilde{v}(n,i)\right\} \; ; \; i = 1, \ldots, M \; ; \; n = 0, 1, \ldots$$

from which we conclude that

$$P_e(M,\{t_j\},\{\kappa_j\},T) \geq e^{-(\sum_j t_j\kappa_j)AT}\sum_{n=0}^{\infty}\frac{[(\sum_j t_j\kappa_j)AT]^n}{n!}\left[1-\frac{1}{M\tilde{v}(n,i)}\right]_+ . \tag{23}$$

Let $f(r) \stackrel{\triangle}{=} \left(\sum_j t_j\left[\frac{\kappa_j}{\sum_j t_j\kappa_j}\right]^{\frac{1}{r}}\right)^r$, noting that $[t_0, t_1]$ is a probability vector then by [30, Appendix 3A inequality (h)] $f(r_1) \geq f(r_2)$ for $0 < r_1 < r_2$ thus $f(r)$ is strictly decreasing. Next let $0 < \theta < 1$ and define $u = \theta r_1 + (1-\theta)r_2$ then by [25, Appendix 5B]

$$f(u) \leq [f(r_1)]^{\theta}[f(r_2)]^{1-\theta}$$

thus $f(r)$ is log-convex on $[0,\infty)$ and hence it is convex $\cup$ [32, pp. 18].

For α defined as

$$\alpha \stackrel{\triangle}{=} \tilde{v}(n,i)^{\frac{1}{n}} = \left[\sum_j t_j\left(\frac{\kappa_j}{\sum_j t_j\kappa_j}\right)^{\frac{1}{1+\rho}}\right]^{\frac{1+\rho}{\rho}},$$

we have $\ln\frac{1}{\alpha} = -\frac{1}{\rho}\ln f(1+\rho)$, consequently it follows that

$$\begin{aligned}\lim_{\rho\to 0}\ln\frac{1}{\alpha} &= -\frac{\partial \ln f(1+\rho)}{\partial\rho}\Big|_{\rho=0} = \frac{f'(1+\rho)}{f(1+\rho)}\Big|_{\rho=0} = f'(1+\rho)\Big|_{\rho=0} \\ &= -\ln\sum_j t_j\kappa_j + \frac{1}{\sum_j t_j\kappa_j}\sum_j t_j\kappa_j\ln\kappa_j \stackrel{\triangle}{=} \vartheta .\end{aligned} \tag{24}$$

Using the α definition we can write (23) as

$$P_e(M,\sum_j t_j\kappa_j,T) \geq e^{-(\sum_j t_j\kappa_j)AT}\sum_{n=0}^{\infty}\frac{[(\sum_j t_j\kappa_j)AT]^n}{n!}\left[1-\frac{1}{M\alpha^n}\right]_+ \tag{25}$$

Thus α reflects the effective pulse energy that in a non-dark-current situation would have generated the same volume $\tilde{v}(n,i)$ that is being experienced in our scenario (compare with [21, Appendix II]). In particular, since we analyze rates close to and above capacity we need the limiting value of $\ln(1/\alpha)$ at capacity (defined above as ϑ).

In [22, section VI Lemma 2] it is shown that the following inequality holds

$$e^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!} \geq \Phi\left(\frac{n-\lambda}{\sqrt{\lambda}}\right), \tag{26}$$

where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-u^2/2}du$.

Combining (25) and (26) we conclude that [22]

$$P_e(M,\sum_j t_j\kappa_j,T) \geq \max_n\left\{\left(1-\frac{1}{M\alpha^n}\right)\Phi\left(\frac{n-\sum_j t_j\kappa_j AT}{\sqrt{n}}\right)\right\} \tag{27}$$

We assume now that R is greater than capacity, i.e. $\ln M = ATf_1(p,q)+a\sqrt{AT} = \sum_j t_j\kappa_j\vartheta AT+a\sqrt{AT}$ where a is an arbitrary positive constant. Since n is defined as the largest integer that satisfies $M\alpha^n \approx 1$ we choose n as

$$n = \frac{\sum_j t_j\kappa_j\vartheta AT + b\sqrt{AT}}{\ln(1/\alpha)|_{\rho=0}} = \frac{\sum_j t_j\kappa_j\vartheta AT + b\sqrt{AT}}{\vartheta}, \quad b < a$$

and get from (27)

$$P_e(M, \sum_j t_j \kappa_j, T) \geq (1 - e^{-(a-b)\sqrt{AT}}) \Phi\left(\frac{b}{\sqrt{\sum_j t_j \kappa_j \vartheta^2 + b\vartheta (AT)^{-1/2}}} \right) . \tag{28}$$

Following [22, section VI] we choose now $b = a - (AT)^{-1/4}$ and get from (28)

$$\liminf_{T \to \infty} P_e(M, \sum_j t_j \kappa_j, T) \geq \Phi\left(\frac{a}{\sqrt{\sum_j t_j \kappa_j \vartheta^2}} \right) , \tag{29}$$

proving that the average error probability is bounded away from zero for rates exceeding capacity.

IV. Proof of the Two-User Result

To evaluate the capacity in this case we use (again) the DMC decomposition and Gubner's inner and outer bounds for the two-user capacity region [10], which are shown to coincide in the case at hand. Then using the code construction of [21] and a suitable "minimum-distance" decoding rule we evaluate an error exponent for this decoder and verify that the capacity as claimed is attainable via this decoder.

A. *Capacity Evaluation via the DMC Decomposition*

Subject to the assumptions a, b in section II the channel reduces to a user-user binary-input binary-output discrete memoreless multiple-access arbitrarily varying channel with transition probability $W(y|x_1, x_2, s_1)$ given by

$$\begin{aligned}
W(1|0,0,0) &- s_0\Delta + O(\Delta^2) , \\
W(1|0,0,1) &- W(1|0,1,0) = W(1|1,0,0) = (1 + s_0)\Delta + O(\Delta^2) , \\
W(1|1,1,0) &- W(1|1,0,1) = W(1|0,1,1) = (2 + s_0)\Delta + O(\Delta^2) \\
W(1|1,1,1) &- (3 + s_0)\Delta + O(\Delta^2)
\end{aligned} \tag{30}$$

Again, since the inputs' and state alphabets are finite the results of [9, 10] are applicable.

For a state distribution $Q = (1 - q, q)$, consider the channel $W_Q \ : \ \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}$

$$\begin{aligned}
W_Q(1|0,0) & \quad (1-q)s_0\Delta + q(1+s_0)\Delta + O(\Delta^2) = \Delta(s_0 + q) + O(\Delta^2) \\
W_Q(1|1,0) &- W_Q(1|0,1) = (1-q)(1+s_0)\Delta + q(2+s_0)\Delta + O(\Delta^2) = \Delta(s_0 + q + 1) + O(\Delta^2) \\
W_Q(1|1,1) &- (1-q)(2+s_0)\Delta + q(3+s_0)\Delta + O(\Delta^2) = \Delta(s_0 + q + 2) + O(\Delta^2) .
\end{aligned}$$

Hence for input distributions $P_i = (1 - p_i, p_i)$, $i = 1, 2$ the mutual information $I(P_1, P_2, W_Q)$ denoted by $I(p_1, p_2, q)$, equals

$$\begin{aligned}
I(p_1, p_2, q) &= h(\Delta(s_0 + q + p_1 + p_2)) - (1 - p_1)(1 - p_2)h(\Delta(s_0 + q)) \\
&\qquad -(p_1 + p_2 - 2p_1p_2)h(\Delta(s_0 + q + 1)) - p_1p_2h(\Delta(s_0 + q + 2)) \\
&- \Delta f(p_1, p_2, q) .
\end{aligned} \tag{31}$$

Similarly, it can be easily verified that $I(P_1, P_2, W_Q|X_1) \stackrel{\triangle}{=} I(p_1, p_2, q|X_1) = \Delta g(p_1, p_2, q)$. Furthermore, by differentiation and use of Jensen's inequality, it can be verified that both $I(p_1, p_2, q)$ and $I(p_1, p_2, q|X_1)$ are decreasing functions of q, then $\ell(Q) = q$ implies

$$\begin{aligned} \min_{q\leq\Lambda} I(p_1, p_2, q) &= I(p_1, p_2, \Lambda) = f(p_1, p_2, \Lambda) \\ \min_{q\leq\Lambda} I(p_1, p_2, q|X_1) &= I(p_1, p_2, \Lambda|X_1) = g(p_1, p_2, \Lambda) \ . \end{aligned} \tag{32}$$

The random-code capacity as claimed in Theorem 2 follows straightforwardly from (32) via [11].

Turning to deterministic codes we consider the set of channels $\mathcal{U}_{\mathcal{X}_1\mathcal{X}_2}(W) \stackrel{\triangle}{=} U : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{S}$ that satisfy

$$\sum_{s_1} W(y|x_1, x_2, s_1)U(s_1|x_1', x_2') = \sum W(y|x_1', x_2', s_1)U(s_1|x_1, x_2) \ , \text{ for all } x_1, x_1', x_2, x_2', y \ . \tag{33}$$

It suffices to consider (33) with $y = 1$ and $(x_1, x_2) = (0, 0), (1, 0)$ then it is easy to show that if U satisfies (33) then $U(\cdot|0,0) = U(\cdot|1,1) = 0$ while $U(0|0,1) = U(0|1,0) = 1 - U(0|0,1) = 1 - U(0|1,0) = t$, $0 < t < 1$. Since this is not a transition probability, the two-user OR channel (30) is nonsymmetrizable-$\mathcal{X}_1\mathcal{X}_2$. Consequently, $l^W_{\mathcal{X}_1\mathcal{X}_2} = \infty$.

Next we determine $l^W_{\mathcal{X}_1}(p_1)$, namely the minimum of

$$\sum_{x_1, s_1} P_1(x_1)U(s_1|x_1)\ell(s_1) = (1 - p_1)U(1|0) + p_1 U(1|1)$$

over the set of channels $\mathcal{U}_{\mathcal{X}_1}(W) \stackrel{\triangle}{=} U : \mathcal{X}_1 \to \mathcal{S}$ such that

$$\sum W(y|x_1, x_2, s_1)U(s_1|x_1') = \sum_{s_1} W(y|x_1', x_2, s_1)U(s_1|x_1) \ , \text{ for all } x_1, x_1', y \ . \tag{34}$$

It suffices to consider $y = 1, x_1 = 0, x_1' = 1$, then

$$\sum_{s_1} W(1|0, x_2, s_1)U(s_1|1) = \sum_{s_1} W(1|1, x_2, s_1)U(s_1|0) \ ,$$

implies the relations

$$\begin{aligned} s_0 U(0|1) + (1 + s_0)U(1|1) &- (1 + s_0)U(0|0) + (2 + s_0)U(1|0) \\ (1 + s_0)U(0|1) + (2 + s_0)U(1|1) &- (2 + s_0)U(0|0) + (3 + s_0)U(1|0) \end{aligned}$$

which are satisfied when $U(0|0) = U(1|1) = 1$.

Thus $l^W_{\mathcal{X}_1}(p_1) = p_1$, and $l^W_{\mathcal{X}_1} = \max_{p_1} l^W_{\mathcal{X}_1}(p_1) = 1$. Consequently, if $\Lambda < 1$ the two-user capacity region has a non-empty interior [10].

Let $P_2(x_2)$ be any probability distribution on $\mathcal{X}_2$, set $(P_2W)(y|x_1, s_1) \stackrel{\triangle}{=} \sum_{x_2} P_2(x_2)W(y|x_1, x_2, s_1)$ and let $\mathcal{U}_{\mathcal{X}_1}(P_2, W)$ denote the set of channels $U : \mathcal{X}_1 \to \mathcal{S}$ such that

$$\sum (P_2W)(y|x_1, s_1)U(s_1|x_1') = \sum_{s_1} (P_2W)(y|x_1', s_1)U(s_1|x_1) \ , \text{ for all } x_1, x_1', y \ . \tag{35}$$

It can be verified that for the OR channel (30) $\mathcal{U}_{\mathcal{X}_1}(P_2, W) = \mathcal{U}_{\mathcal{X}_1}(W)$. Consequently, by [10, Lemma 5.6 and Theorem 5.8] and [12] the rate region in Theorem 2 is achievable provided that the state constraint satisfies $\Lambda < \min\{l^W_{\mathcal{X}_1}(p_1), l^W_{\mathcal{X}_2}(p_2)\} = \min\{p_1, p_2\}$. Since this region equals also to the upper bound given by [10, Theorem 3.2] this establishes the deterministic-code capacity as claimed in Theorem 2.

B. *Code Construction and Decoding Rule*

We describe a family of two-user codes with parameters T, M_1, M_2 where the code waveforms satisfy the peak and average power constraint p_1, p_2. To this end we set $M_1 = \lceil e^{R_1 T} \rceil$, $M_2 = \lceil e^{R_2 T} \rceil$, $M = M_1 + M_2$. Again it suffices to construct only the supports $b_i = \{t : \lambda_i(t) = A\}$. Start with the case $p_1 = p_2 = p$ and apply Wyner's construction to obtain a simplex system of measurable supports $\{b_1, \ldots, b_M\} \subseteq [0, T]$, then partition arbitrarily that system into two subsets $\{M_1\} = \{b_1, \ldots, b_{M_1}\}$ and $\{M_2\} = \{b_{M_1+1}, \ldots, b_M\}$ of cardinalities M_1, M_2, respectively. In the case $p_1 = p_2 = p$ the subsets $\{M_1\}$ and $\{M_2\}$ form the first and the second user codebooks, respectively. Denote a codebook of this form by $\mathcal{C}(M_1, M_2, p)$.

When $p_1 \neq p_2$ assume, without loss of generality, that $p_1 < p_2$. Then we generate first $\mathcal{C}(M_1, M_2, p_2)$ and modify the simplex system $\{M_1\}$ replacing it by another simplex system $\{M_1^*\}$ corresponding to the parameter p_1. Recall that the interval $[0, T]$ was partitioned on small subintervals Δ_j where each of the M supports $\{b_1, \ldots, b_M\}$ could only contain complete (or didn't contain any) subintervals Δ_j. Denote $p = p_1/p_2 < 1$. Consider one subinterval Δ_j and construct on it a new simplex system of supports $\{b_{j,l}\} \subset \Delta_j$, corresponding to the parameter p. Now for any $b_i \in \{M_1\}$ such that $\Delta_j \subset b_i$ replace that Δ_j with *any arbitrary* "refinement" $b_{j,l}$. This procedure is done for all Δ_j and all $b_i \in \{M_1\}$. As a result we get a pair of simplex systems, corresponding to the parameters p_1, p_2, with the following distance properties [21].

$$
\begin{aligned}
\frac{1}{T}\mu\{t : \lambda_m(t) = A\} &\approx p_1 && m \in \{M_1\} \\
\frac{1}{T}\mu\{t : \lambda_m(t) = A\} &\approx p_2 && m \in \{M_2\} \\
\frac{1}{T}\mu\{t : \lambda_{m_1}(t) = A \ , \ \lambda_{m_2}(t) = 0\} &\approx p_1(1 - p_2) \ , && m_1 \in \{M_1\}, m_2 \in \{M_2\} \\
\frac{1}{T}\mu\{t : \lambda_{m_1}(t) = 0 \ , \ \lambda_{m_2}(t) = A\} &\approx (1 - p_1)p_2 \ , && m_1 \in \{M_1\}, m_2 \in \{M_2\} \ .
\end{aligned}
$$

Decoder : To define the decoder mapping D for the two-user code two cases should be considered.

Case I: Joint detection of both messages.

Any hypothesis $m = (m_1, m_2)$ defines three independent Poisson processes; the first with transmission rate $\zeta_m(t) = A + \lambda_0 + \lambda_{s_1}(t)$ where only one codebook - m_1 or m_2 - is "active", the second with transmission rate $\zeta_m(t) = 2A + \lambda_0 + \lambda_{s_1}(t)$ where both - m_1 and m_2 - are active, and the third with transmission rate $\zeta_m(t) = \lambda_0 + \lambda_{s_1}(t)$ where none of the users is active.

For $m = (m_1, m_2), 1 \leq m_1 \leq M_1, M_1 + 1 \leq m_2 \leq M_1 + M_2$ let (cf. Figure 2)

$$
\begin{aligned}
S_m^{10} &\stackrel{\Delta}{=} b_{m_1} \cap b_{m_2}^c \\
S_m^{01} &\stackrel{\Delta}{=} b_{m_1}^c \cap b_{m_2} \\
S_m^{11} &\stackrel{\Delta}{=} b_{m_1} \cap b_{m_2}
\end{aligned}
$$

$$S_m^{00} \;\triangleq\; b_{m_1}^c \cap b_{m_2}^c \; .$$

The decoder observes ν_0^T, computes the sufficient statistics vector $(\psi_m^0, \psi_m^1, \psi_m^\phi)$ defined as

$$\begin{aligned} & - \int_{S_m^{10} \cup S_m^{01}} d\nu(t) = \{\text{number of arrivals in } S_m^{10} \cup S_m^{01}\} \\ \psi_m^1 \quad & \int_{S_m^{11}} d\nu(t) = \{\text{number of arrivals in } S_m^{11}\} \; , \\ \psi_m^\phi \quad & - \quad n - (\psi_m^0 + \psi_m^1) = \{\text{number of arrivals in } S_m^{00}\}, \end{aligned}$$

and $D(\nu_0^T) = m \triangleq (m_1, m_2)$ iff

- $\psi_m^1 + \psi_m^0 \geq \psi_{m'}^1 + \psi_{m'}^0$.
- No m' exists such that $\psi_{m'}^1 \geq \psi_m^1$ and $\psi_{m'}^0 \geq \psi_m^0$.

That is, given the observation ν_0^T the decoder decodes the message m that is "nearest" to the observation ν_0^T and not dominated by any other message m' with respect to both projections ψ_m^1 and ψ_m^0. The reader familiar with the class of β-decoders defined in [3, 4] (see also the tutorial [14, section IV.6]) realizes that this decoding rule falls within this category.

Consequently,

$$\begin{aligned} P_{em} \;&\triangleq\; \Pr\{D(\nu_0^T) = m' | m\} \\ &< \; \Pr\Big(\bigcup_{m' \neq m} \{\psi_{m'}^1 \geq \psi_m^1\} \cap \{\psi_{m'}^0 \geq \psi_m^0\} | m \Big) \triangleq \Pr\Big(\bigcup_{m' \neq m} \mathbf{E}_{m'} | m \Big) \; , \end{aligned} \tag{36}$$

where $\mathbf{E}_{m'}$ denotes the event $\{\psi_{m'}^1 \geq \psi_m^1\} \cap \{\psi_{m'}^0 \geq \psi_m^0\}$.

Case II: Detection of one message while the other is known.

Assume that the decoder knows λ_{m_1} and wishes to decode λ_{m_2}. In this case the decoder treats the known message λ_{m_1} as noise, on the interval where it is active, while decoding λ_{m_2}. Consequently the decoding interval is split into two regions. For any $M_1 + 1 \leq m_2 \leq M_1 + M_2$ let

$$\begin{aligned} S_m^0 \;&\triangleq\; \{t \in [0,T] : \lambda_{m_1} = 0 \text{ and } \lambda_{m_2} = A\} \\ S_m^1 \;& \quad\; \{t \in [0,T] : \lambda_{m_1} = A \text{ and } \lambda_{m_2} = A\} \end{aligned}$$

The decoder observes ν_0^T and computes

$$\begin{aligned} & \triangleq \int_{S_m^0} d\nu(t) = \{\text{number of arrivals in } S_m^0\} \\ & \triangleq \int_{S_m^1} d\nu(t) = \{\text{number of arrivals in } S_m^1\} \end{aligned}$$

Then $D(\nu_0^T) = m_2^*$ if $\psi_{m_2}^0 \leq \psi_{m_2^*}^0$ and $\psi_{m_2}^1 \leq \psi_{m_2^*}^1$ $\forall m_2 \neq m_2^*$ - i.e. the decoder favors the message m_2^* which maximizes the number of arrivals on both intervals.

C. *Error probability for detection of both users*

Let $m = (m_1, m_2)$ be the transmitted message, $m' = (m'_1, m'_2) \neq m$ any other message. We assume the support set $S_m = S_{m_1} \cup S_{m_2}$ of $\left(\lambda_{m_1}(\cdot), \lambda_{m_2}(\cdot)\right)$ as shown in Fig. 2. This implies that the number of arrivals in $[0, T]$ is the sum of three independent Poisson distributed random variables, $W_B = \nu(\tau_2) - \nu(\tau_1)$, $W_{AC} = \nu(\tau_1) + \nu(\tau_3) - \nu(\tau_2)$ and $W_D = \nu(T) - \nu(\tau_3)$, with parameters Λ_2, Λ_1 and Λ_0 respectively, where

$$\begin{aligned} \Lambda_2 &\quad (2A + qA + \lambda_0)p_1p_2T = (2 + s_0 + q)p_1p_2AT \\ \Lambda_1 &- (A + qA + \lambda_0)(p_1 + p_2 - 2p_1p_2)T = (1 + s_0 + q)(p_1 + p_2 - 2p_1p_2)AT \\ \Lambda_0 &\quad (qA + \lambda_0)(1 - p_1)(1 - p_2)T = (s_0 + q)(1 - p_1)(1 - p_2)AT \end{aligned} \tag{37}$$

Thus $\nu(T)$ is Poisson distributed with parameter $\Lambda = \Lambda_0 + \Lambda_1 + \Lambda_2 = (s_0 + q + p_1 + p_2)AT$. Furthermore, given that $\nu(T) = n$, W_B, W_{AC} and W_D have the extended Bernoulli (multinomial) distribution

$$\Pr\left\{W_{AC} = n_1, W_B = n_2, W_D = n - n_1 - n_2 | \lambda_{m_1}, \lambda_{m_2}, \nu(T) = n\right\} = \\ \binom{n}{n_1}\binom{n - n_1}{n_2}\pi_1^{n_1}\pi_2^{n_2}(1 - \pi_1 - \pi_2)^{n - n_1 - n_2} \triangleq P(n_1, n_2, n) \tag{38}$$

where

$$\pi_i = \frac{\Lambda_i}{\Lambda_0 + \Lambda_1 + \Lambda_2} \qquad i = 1, 2 \ .$$

Thus the joint probability

$$\Pr\left\{W_{AC} = n_1, W_B = n_2, W_D = n - n_1 - n_2, \nu(T) = n | \lambda_{m_1}, \lambda_{m_2}\right\} = \\ \frac{e^{-\Lambda}\Lambda^n}{n!}\binom{n}{n_1}\binom{n - n_1}{n_2}\pi_1^{n_1}\pi_2^{n_2}(1 - \pi_1 - \pi_2)^{n - n_1 - n_2} \triangleq Q(n_1, n_2, n) \tag{39}$$

The code waveforms $\left(\lambda_{m'_1}(\cdot), \lambda_{m'_2}(\cdot)\right)$ define the corresponding subregions A'_0, B'_0 and C'_0 defined as

$$\begin{aligned} A'_0 &- b_{m'_1} \cap b^c_{m'_2} \qquad \mu(A'_0) = p_1(1 - p_2) \\ C'_0 &- b^c_{m'_1} \cap b_{m'_2} \qquad \mu(C'_0) = p_2(1 - p_1) \\ B'_0 &- b_{m'_1} \cap b_{m'_2} \qquad \mu(B'_0) = p_1p_2 \end{aligned}$$

Given that $W_B = n_2$ we are interested in how these arrivals are distributed on $A'_0 + C'_0$, B'_0 and $(A'_0 + C'_0 + B'_0)^c$. Let W_{10}, W_{11} and $W_{1\phi}$ denote the number of arrivals on $(A'_0 + C'_0) \cap B_0$, on $B'_0 \cap B_0$ and on $(A'_0 + C'_0 + B'_0)^c \cap B_0$ then $(W_{10}, W_{11}, W_{1\phi})$ are jointly distributed according to

$$\Pr\left\{W_{10} = k_0, W_{11} = k_1, W_{1\phi} = n_2 - k_0 - k_1\right\} = \binom{n_2}{k_0}\binom{n_2 - k_0}{k_1}t_1^{k_0}t_2^{k_1}(1 - t_1 - t_2)^{n_2 - k_0 - k_1} \ , \tag{40}$$

with $t_1 = \mu(A'_0 + C'_0) = (p_1 + p_2 - 2p_1p_2)$, $t_2 = \mu(B'_0) = p_1p_2$.

Similarly, the n_1 arrivals on W_{AC} and n_0 on W_D are distributed on $A_0'+C_0'$, B_0' and $(A_0'+C_0'+B_0')^c$ according to

$$\Pr\Big\{W_{00}=k_0, W_{01}=k_1, W_{0\phi}=n_1-k_0-k_1\Big\} \quad \binom{n_1}{k_0}\binom{n_1-k_0}{k_1} t_1^{k_0} t_2^{k_1}(1-t_1-t_2)^{n_1-k_0-k_1}$$
$$\Pr\Big\{W_{\phi 0}=k_0, W_{\phi 1}=k_1, W_{\phi\phi}=n_0-k_0-k_1\Big\} - \binom{n_0}{k_0}\binom{n_0-k_0}{k_1} t_1^{k_0} t_2^{k_1}(1-t_1-t_2)^{n_0-k_0-k_1} \quad (41)$$

Furthermore, we will also need the conditional distributions of the n_2 arrivals on W_B and n_0 arrivals on W_D given that the n_1 arrivals on W_{AC} are split to : k_{00} on $A_0'+C_0'$, k_{10} on B_0' and $k_{\phi 0}$ on $(A_0'+C_0'+B_0')^c$, namely

$$\Pr\Big\{W_{11}=k_1, W_{1\phi}=n_2-k_1-k_{10} | W_{10}=k_{10}\Big\} = \binom{n_2-k_{10}}{k_1}\frac{t_2^{k_1}(1-t_1-t_2)^{n_2-k_{10}-k_1}}{(1-t_1)^{n_2-k_{10}}}$$
$$\Pr\Big\{W_{01}=k_1, W_{0\phi}=n_1-k_1-k_{00} | W_{00}=k_{00}\Big\} = \binom{n_1-k_{00}}{k_1}\frac{t_2^{k_1}(1-t_1-t_2)^{n_1-k_{00}-k_1}}{(1-t_1)^{n_1-k_{00}}}$$
$$\Pr\Big\{W_{\phi 1}=k_1, W_{\phi\phi}=n_0-k_1-k_{\phi 0} | W_{\phi 0}=k_{\phi 0}\Big\} = \binom{n_0-k_{\phi 0}}{k_1}\frac{t_2^{k_1}(1-t_1-t_2)^{n_0-k_{\phi 0}-k_1}}{(1-t_1)^{n_0-k_{\phi 0}}} \quad (42)$$

We define now the following two events $\mathcal{A}_1$ and $\mathcal{A}_2$ as

$$\mathcal{A}_1 \stackrel{\triangle}{=} \Big(W_{00}+W_{10}+W_{\phi 0}-n_1 \geq 0 | \lambda_{m_1}, \lambda_{m_2}, n_2, n_1, n\Big)$$
$$\mathcal{A}_2 \stackrel{\triangle}{=} \Big(W_{01}+W_{11}+W_{\phi 1}-n_2 \geq 0 | \lambda_{m_1}, \lambda_{m_2}, n_2, n_1, n, W_{00}=k_{00}, W_{10}=k_{10}, W_{\phi 0}=k_{\phi 0}\Big) . \quad (43)$$

Notice that $\mathcal{A}_1 \cap \mathcal{A}_2$ actually determines the sought for event $\mathbf{E}_{\mathbf{m}'}$.

From (40), (41) and (42)

$$E\Big(W_{00}+W_{10}+W_{\phi 0}-n_1 | \lambda_{m_1}, \lambda_{m_2}, n_1, n_2, n\Big) - nt_1 - n_1$$
$$E\Big(W_{01}+W_{11}+W_{\phi 1}-n_2 | \lambda_{m_1}, \lambda_{m_2}, n_1, n_2, n, W_{00}, W_{10}, W_{\phi 0}\Big) - (n-n_1)\frac{t_2}{1-t_1} - n_2$$

If the above expectations are negative, the probabilities $\Pr(\mathcal{A}_1)$ and $\Pr(\mathcal{A}_2)$ will be small. Hence we define the following sets

$$A_0 - \Big\{(n_1, n) : 0 \leq n_1 \leq n \ , \ nt_1 - n_1 < 0\Big\}$$
$$A_1 - \Big\{(n_2, n) : 0 \leq n_2 \leq n - n_1 \ , \ (n-n_1)t_2 - n_2(1-t_1) < 0\Big\} .$$

Next we have

$$\Pr\Big(\mathcal{A}_1\Big) \leq E\Big(e^{\tau(W_{00}+W_{10}+W_{\phi 0}-n_1)} | \lambda_{m_1}, \lambda_{m_2}, n_1, n_2, n\Big)$$

Using the distributions in (40) and (41), we obtain

$$\Pr\Big(\mathcal{A}_1\Big) \leq \exp\{\gamma_1(\tau)\} = \exp\{n \ln(1-t_1+t_1 e^{\tau}) - n_1\tau\} .$$

$\gamma_1(\tau)$ is minimized when $e^\tau = \frac{n_1(1-t_1)}{(n-n_1)t_1}$, - i.e. $\tau = \ln \frac{n_1(1-t_1)}{(n-n_1)t_1} \geq 0$ if $(n_1, n) \in A_0$. Substituting this value we get

$$\Pr\Big(\mathcal{A}_1\Big) \;\leq\; \frac{n^n(1-t_1)^{n-n_1}t_1^{n_1}}{(n-n_1)^{n-n_1}n_1^{n_1}} \triangleq \Gamma_1(n_1, n_2, n) \; . \tag{44}$$

In a similar way

$$\Pr\Big(\mathcal{A}_2\Big) \leq E\Big(e^{\tau(W_{01}+W_{11}+W_{\phi 1}-n_2)} | \lambda_{m_1}, \lambda_{m_2}, n_1, n_2, n, W_{00}, W_{10}, W_{\phi 0}\Big)$$

Using the conditional distributions in (42) we obtain

$$\Pr\Big(\mathcal{A}_2\Big) \leq \exp\{\gamma_2(\tau)\} = \exp\{(n-n_1)[\ln(t_2 e^\tau + (1-t_1-t_2)) - \ln(1-t_1)] - n_2\tau\} \; .$$

$\gamma_2(\tau)$ is minimized when $e^\tau = \frac{n_2(1-t_1-t_2)}{(n-n_1-n_2)t_2}$, - i.e. $\tau = \ln \frac{n_2(1-t_1-t_2)}{(n-n_1-n_2)t_2} \geq 0$ if $(n_1, n_2, n) \in A_1$. Substituting this value we get

$$\Pr\Big(\mathcal{A}_2\Big) \leq \frac{(n-n_1)^{n-n_1}}{(1-t_1)^{n-n_1}} \cdot \frac{t_2^{n_2}(1-t_1-t_2)^{n-n_1-n_2}}{n_2^{n_2}(n-n_1-n_2)^{n-n_1-n_2}} \triangleq \Gamma_2(n_1, n_2, n) \tag{45}$$

Combining (44) and (45) we conclude that

$$\Pr\Big(\mathcal{A}_1\Big) \cdot \Pr\Big(\mathcal{A}_2\Big) \leq \frac{n^n t_1^{n_1} t_2^{n_2}(1-t_1-t_2)^{n-n_1-n_2}}{n_1^{n_1} n_2^{n_2}(n-n_1-n_2)^{n-n_1-n_2}} \triangleq \Gamma(n_1, n_2, n) \; .$$

In summary we have

$$\begin{aligned}
P_{em} \;&\leq\; \Pr\Big(\bigcup_{m' \neq m} \mathbf{E}_{\mathbf{m}'} | \lambda_{m_1}, \lambda_{m_2}\Big) \\
&\quad \sum_{n_1, n_2, n} Q(n_1, n_2, n) \Pr\Big(\bigcup_{m' \neq m} \mathbf{E}_{\mathbf{m}'} | \lambda_{m_1}, \lambda_{m_2}, W_{AC} = n_1, W_B = n_2, \nu(T) = n\Big) \\
&< \sum_{(n_1, n_2, n) \in A_0 \cap A_1} Q(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma(n_1, n_2, n)\Big]^\rho \\
&\quad + \sum_{(n_1, n_2, n) \in A_0 \cap A_1^c} Q(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma_1(n_1, n_2, n)\Big]^\rho \\
&\quad + \sum_{(n_1, n_2, n) \in A_0^c \cap A_1} Q(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma_2(n_1, n_2, n)\Big]^\rho \\
&\quad + \sum_{(n_1, n_2, n) \in A_0^c \cap A_1^c} Q(n_1, n_2, n) \\
&\triangleq\; P_{em}(1) + P_{em}(2) + P_{em}(3) + P_{em}(4)
\end{aligned} \tag{46}$$

Consider the term

$$\begin{aligned}
P_{em}(1) \;&-\; \sum_{(n_1, n_2, n) \in A_0 \cap A_1} Q(n_1, n_2, n) \Big[\sum_{(m_1', m_2') \neq (m_1, m_2)} \Gamma(n_1, n_2, n)\Big]^\rho \\
&<\; M_1^\rho M_2^\rho \sum_{n=1}^{\infty} \sum_{n_1=0}^{n} \sum_{n_2=0}^{n-n_1} Q(n_1, n_2, n) \Big[\Gamma(n_1, n_2, n)\Big]^\rho
\end{aligned}$$

Now for a given n,

$$\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1} Q(n_1,n_2,n)\Gamma^\rho(n_1,n_2,n) =$$
$$\frac{e^{-\Lambda}\Lambda^n}{n!}\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1}\binom{n}{n_1}\binom{n-n_1}{n_2}\pi_1^{n_1}\pi_2^{n_2}(1-\pi_1-\pi_2)^{n-n_1-n_2}\left[\frac{t_1^{n_1}t_2^{n_2}(1-t_1-t_2)^{n-n_1-n_2}n^n}{n_1^{n_1}n_2^{n_2}(n-n_1-n_2)^{n-n_1-n_2}}\right]^\rho$$

Let $\xi_1 = n_1/n$, $\xi_2 = n_2/n$ so that

$$\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1} Q(n_1,n_2,n)\Gamma^\rho(n_1,n_2,n) =$$
$$\frac{e^{-\Lambda}\Lambda^n}{n!}\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1}\binom{n}{\xi_1 n}\binom{n-\xi_1 n}{\xi_2 n}\pi_1^{\xi_1 n}\pi_2^{\xi_2 n}(1-\pi_1-\pi_2)^{(1-\xi_1-\xi_2)n}\left[\frac{t_1^{\xi_1}t_2^{\xi_2}(1-t_1-t_2)^{1-\xi_1-\xi_2}}{\xi_1^{\xi_1}\xi_2^{\xi_2}(1-\xi_1-\xi_2)^{1-\xi_1-\xi_2}}\right]^{\rho n} \quad (47)$$

Using the fact (can be derived in the same way as Example 12.1.3 in [31])

$$\begin{aligned}\binom{n}{\xi_1 n}\binom{n-\xi_1 n}{\xi_2 n} &< e^{nH(\xi_1,\xi_2,1-\xi_1-\xi_2)}\\ &= \exp\{-n[\xi_1\ln\xi_1+\xi_2\ln\xi_2+(1-\xi_1-\xi_2)\ln(1-\xi_1-\xi_2)]\}\ ,\end{aligned}$$

(47) becomes

$$\begin{aligned}&\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1} Q(n_1,n_2,n)\Gamma^\rho(n_1,n_2,n)\\ &\le \frac{e^{-\Lambda}\Lambda^n}{n!}\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1}\exp\left\{n\left[\xi_1\ln\frac{t_1^\rho\pi_1}{\xi_1^{1+\rho}}+\xi_2\ln\frac{t_2^\rho\pi_2}{\xi_2^{1+\rho}}+(1-\xi_1-\xi_2)\ln\frac{(1-t_1-t_2)^\rho(1-\pi_1-\pi_2)}{(1-\xi_1-\xi_2)^{1+\rho}}\right]\right\}\\ &\le \frac{e^{-\Lambda}\Lambda^n}{n!}n(n-1)\exp\left\{n(1+\rho)\max_{0\le\xi_1+\xi_2\le1}\varphi(\xi_1,\xi_2)\right\}\end{aligned}$$

where

$$\varphi(\xi_1,\xi_2)\triangleq\xi_1\ln\frac{\alpha}{\xi_1}+\xi_2\ln\frac{\beta}{\xi_2}+(1-\xi_1-\xi_2)\ln\frac{\gamma}{1-\xi_1-\xi_2}\ , \quad (48)$$

and we've used the notation $\alpha = (t_1^\rho\pi_1)^{\frac{1}{1+\rho}}$, $\beta = (t_2^\rho\pi_2)^{\frac{1}{1+\rho}}$, $\gamma = [(1-t_1-t_2)^\rho(1-\pi_1-\pi_2)]^{\frac{1}{1+\rho}}$. The function $g(\xi_1,\xi_2)$ is a concave function in ξ_1 and ξ_2 and is maximized for $\xi_1 = \alpha/(\alpha+\beta+\gamma)$ and $\xi_2 = \beta/(\alpha+\beta+\gamma)$ and the maximum is $\ln(\alpha+\beta+\gamma)$, this yields

$$\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1} Q(n_1,n_2,n)\Gamma^\rho(n_1,n_2,n)\le\frac{e^{-\Lambda}\Lambda^n}{(n-2)!}(\alpha+\beta+\gamma)^{n(1+\rho)}$$

from which we get

$$\begin{aligned}P_{em}(1) &< M_1^\rho M_2^\rho\sum_{n=1}^{\infty}\frac{e^{-\Lambda}\Lambda^n}{(n-2)!}(\alpha+\beta+\gamma)^{n(1+\rho)}\\ &= M_1^\rho M_2^\rho\Lambda^2(\alpha+\beta+\gamma)^{2(1+\rho)}\cdot\sum_{n=1}^{\infty}e^{-\Lambda}\frac{(\alpha+\beta+\gamma)^{(1+\rho)(n-2)}\Lambda^{n-2}}{(n-2)!}\\ &\quad M_1^\rho M_2^\rho\Lambda^2(\alpha+\beta+\gamma)^{2(1+\rho)}\exp\left\{-\Lambda+\Lambda(\alpha+\beta+\gamma)^{(1+\rho)}\right\}\end{aligned} \quad (49)$$

Noting that

$$\begin{aligned}
\left(t_1^\rho \pi_1\right)^{\frac{1}{1+\rho}} &= (p_1+p_2-2p_1p_2)\left(\frac{1+s_0+q}{s_0+q+p_1+p_2}\right)^{\frac{1}{1+\rho}} \\
\left(t_2^\rho \pi_2\right)^{\frac{1}{1+\rho}} &= p_1p_2\left(\frac{2+s_0+q}{s_0+q+p_1+p_2}\right)^{\frac{1}{1+\rho}} \\
\left((1-t_1-t_2)^\rho(1-\pi_1-\pi_2)\right)^{\frac{1}{1+\rho}} &= (1-p_1)(1-p_2)\left(\frac{s_0+q}{s_0+q+p_1+p_2}\right)^{\frac{1}{1+\rho}}
\end{aligned}$$

the exponent in the right member of (49) is

$$\begin{aligned}
&-\Lambda+\Lambda(\alpha+\beta+\gamma)^{(1+\rho)} \\
&\quad -TA\Big\{(s_0+q+p_1+p_2)-\Big[(p_1+p_2-2p_1p_2)(1+s_0+q)^{\frac{1}{1+\rho}} \\
&\qquad +p_1p_2(2+s_0+q)^{\frac{1}{1+\rho}}+(1-p_1)(1-p_2)(s_0+q)^{\frac{1}{1+\rho}}\Big]^{1+\rho}\Big\} \\
&\triangleq -TAE_{12}(\rho,p_1,p_2,q)
\end{aligned} \tag{50}$$

In Appendix II we show that the terms $P_{em}(i)$, $i=2,3$ are bounded by $B_i e^{-T(A\hat{E}_i-\rho R)}$ where $\hat{E}_i \geq E_{12}$. Furthermore, it is also shown there that $P_{em}(4) \leq e^{-\hat{E}_4 T}$ and $\hat{E}_4 \geq \max_{0\leq\rho\leq 1} E_{12}(\rho,p_1,p_2,q) = E_{12}(1,p_1,p_2,q)$. Thus the average error probability of our "minimum distance" decoder satisfies

$$P_e = \frac{1}{M_1M_2}\sum_{m_1=1}^{M_1}\sum_{m_2=M_1+1}^{M_1+M_2} P_{em} \leq \exp\Big\{-T\Big[AE_{12}(\rho,p_1,p_2,q)-\rho R\Big]+O(T)\Big\} \tag{51}$$

which establishes the joint decoding error exponent lower bound. Since

$$\left.\frac{\partial E_{12}(\rho,p_1,p_2,q)}{\partial\rho}\right|_{\rho=0} = f(p_1,p_2,q)$$

the rate-sum capacity as claimed in Theorem 2 is achievable by this code and decoding rule.

D. *Error probability for detection of one user*

In this case it is assumed that the decoder is informed on the code waveform λ_{m_1}. Consequently, the decoding of λ_{m_2} is split into two intervals; T_1 the interval over which $\lambda_{m_1}=0$ and T_2 the interval over which $\lambda_{m_1}=A$. As each of these decoding tasks is a *single-user* decoding task we can apply the results of section II to evaluate the corresponding error exponent.

On T_1 the decoder decodes λ_{m_2} with background noise of $\lambda_0+\lambda_{s_1}(t)$ hence the probability of erroneous decision on this interval is upper bounded by

$$P_e(1) \leq \exp\Big\{-T_1A\Big[(s_0+q+p_2)-\Big((1-p_2)(s_0+q)^{\frac{1}{1+\rho}}+p_2(s_0+q+1)^{\frac{1}{1+\rho}}\Big)^{1+\rho}\Big]+T_1\rho R_2\Big\} \tag{52}$$

On T_2 the decoder decodes λ_{m_2} with background noise of $\lambda_0+A+\lambda_{s_1}(t)$ hence the average probability of error on this interval is upper bounded by

$$P_e(2) \leq \exp\Big\{-T_2A\Big[(s_0+q+p_2+1)-\Big((1-p_2)(s_0+q+1)^{\frac{1}{1+\rho}}+p_2(s_0+q+2)^{\frac{1}{1+\rho}}\Big)^{1+\rho}\Big]+T_2\rho R_2\Big\} \tag{53}$$

Substituting $T_1 = (1-p_1)T$ and $T_2 = p_1 T$ and recalling that the decoder will make an error if and only if on both intervals its decision is in favor of $m_2' \neq m_2$ - i.e. $P_{em} = P_e(1) \cdot P_e(2)$, we get the error exponent for decoding one user while the other is known. Let this exponent be defined as $E_{11}(\rho, p_1, p_2, q)$, i.e. $P_e(1) \cdot P_e(2) \leq \exp\{-TAE_{11}(\rho, p_1, p_2, q) + T\rho R_2\}$.

Since

$$\left.\frac{\partial E_{11}(\rho, p_1, p_2, q)}{\partial \rho}\right|_{\rho=0} = g(p_1, p_2, q)$$

the capacity region as claimed in Theorem 2 is achievable by our code and decoding rules.

V. Summary and Conclusions

We study the capacity of the Poisson single-user and two-user (multiple-access) arbitrarily varying channel, subject to input and state constraints. Using a DMC decomposition for the non-bandwidth-limited Poisson channel in combination with the general results of Csiszár-Narayan (for the single-user case) and Gubner (for the two-user case) the deterministic-code capacities of these channels have been obtained. An essential result due to Frey, which determines the optimal jamming strategy in the Poisson regime for the case at hand, has been used to obtain a converse.

We've also considered decoding rules that attain the aforementioned capacities. Specifically, in the single-user case Wyner's [19] minimum distance decoder is shown to have a positive error exponent for all rates not exceeding capacity. In the two-user case we suggest a minimum distance decoder that belongs to the class of β-decoders introduced in [3, 4] for deterministic codes on the (single-user) AVC. This decoder is shown to exhibit a positive error exponent everywhere inside the two-user deterministic-code capacity region.

Appendix I

Proof of Theorem 1; the cases $B > A$ and $B < A$.

Consider the case $B > A$ while $1/T \int \lambda_{s_1}(t)dt = \Lambda A$. Using simple function approximation and passage to the DMC limit as $\Delta \to 0$ establishes the optimality of binary signaling for both the encoder and state inputs [20, 22]. A state sequence with constant intensity and average power ΛA can be generated as follows.

The jammer limits its peak to $\tilde{A}$ where $A < \tilde{A} \leq B$ then it forms a refinement partition of the intervals Δ_j, say $\Delta_{j,l}$ and chooses to transmit $\tilde{A}$ on $\bigcup_l \Delta_{j,l} \subset \Delta_j$ such that

$$\left(\bigcup_l \Delta_{j,l}\right)/\Delta_j = \Lambda \frac{A}{\tilde{A}} \leq 1 \, .$$

A decoder that observes ν_0^T on the partition defined by Δ_j cannot distinguish the above state sequence and that which adopts $\tilde{A} = A$ and transmits with peak A on $\bigcup_l \Delta_{j,l}/\Delta_j = \Lambda$. Consequently, Theorem 1 holds for this case.

Consider now the case $B < A$. Here the most the state input can do is transmit with peak B (which is strictly smaller than A) over Δ_j. Let $\beta \stackrel{\triangle}{=} B/A$ and write for this case

$$W(1|0,0) \;\; = \;\; s_0 \Delta$$

$$\begin{array}{ll} W(1|1,0) & (1+s_0)\Delta \\ W(1|0,1) & (s_0+\beta)\Delta \\ W(1|1,1) & (1+s_0+\beta)\Delta \end{array}$$

We consider now the set of channels $U : \mathcal{X} \to \mathcal{S}$ that render the AVC symmetrizable, i.e

$$\sum_{s_1 \in \mathcal{S}} W(y|x, s_1)U(s_1|x') = \sum W(y|x', s_1)U(s_1|x) \ . \tag{54}$$

It suffices to consider (54) with $y = 1, x = 0, x' = 1$, then

$$\sum_{s_1 \in \mathcal{S}} W(1|0, s_1)U(s_1|1) = \sum_{s_1 \in \mathcal{S}} W(1|1, s_1)U(s_1|0) \ ,$$

implies the relation

$$s_0 U(0|1) + (s_0 + \beta)U(1|1) = (1 + s_0)U(0|0) + (1 + s_0 + \beta)U(1|0) \tag{55}$$

The only U that satisfies (55) is $U \equiv 0$, which is not a transition probability, hence the AVC is non-symmetrizable and its deterministic-code capacity equals the random-code capacity. ∘

Appendix II

In this appendix we upper bound each of the terms $P_{em}(i), i = 2 \dots 4$ in (46). To simplify the expressions we use the notation $s \triangleq s_0 + q$.

Consider the term

$$\begin{aligned} P_{em}(2) &= \sum_{(n_1,n_2,n) \in A_0 \cap A_1^c} Q(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma_1(n_1, n_2, n) \Big]^\rho \\ &< \sum_{(n_1,n_2,n) \in A_1^c} Q(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma_1(n_1, n_2, n) \Big]^\rho \\ & \sum_{(n_1,n_2,n)} Q(n_1, n_2, n) \chi(n_1, n_2, n) \Big[\sum_{m' \neq m} \Gamma_1(n_1, n_2, n) \Big]^\rho \\ &< M_1^\rho M_2^\rho \sum_{(n_1,n_2,n)} Q(n_1, n_2, n) e^{\tau[(n-n_1)\frac{t_2}{1-t_1} - n_2]} \Gamma_1^\rho(n_1, n_2, n) \ , \end{aligned}$$

where

$$\chi(n_1, n_2, n) = \begin{cases} 1 & \text{if } (n_1, n_2, n) \in A_1^c \\ 0 & \text{otherwise} \end{cases}$$

and the last inequality follows since $\chi(n_1, n_2, n) \leq \exp\{\tau[(n - n_1)\frac{t_2}{1-t_1} - n_2]\}$.

Now for a given n,

$$\begin{aligned} &\sum_{n_1=0}^{n} \sum_{n_2=0}^{n-n_1} Q(n_1, n_2, n) e^{\tau[(n-n_1)\frac{t_2}{1-t_1} - n_2]} \Gamma_1^\rho(n_1, n_2, n) = \\ &\frac{e^{-\Lambda}\Lambda^n}{n!} \sum_{n_1=0}^{n} \sum_{n_2=0}^{n-n_1} \binom{n}{n_1} \binom{n-n_1}{n_2} \pi_1^{n_1} \left[\pi_2 e^{-\tau(1-\frac{t_2}{1-t_1})}\right]^{n_2} \\ &\cdot \left[(1 - \pi_1 - \pi_2)e^{\tau\frac{t_2}{1-t_1}}\right]^{n-n_1-n_2} \left[\frac{t_1^{n_1}(1-t_1)^{n-n_1} n^n}{(n-n_1)^{n-n_1} n_1^{n_1}}\right]^\rho \end{aligned}$$

To proceed further let us consider first the inner sum

$$\sum_{n_2=0}^{n-n_1} \binom{n-n_1}{n_2} \left[\pi_2 e^{-\tau(1-\frac{t_2}{1-t_1})}\right]^{n_2} \left[(1-\pi_1-\pi_2)e^{\tau\frac{t_2}{1-t_1}}\right]^{n-n_1-n_2}$$

$$= \left[(1-\pi_1-\pi_2)e^{\tau\frac{t_2}{1-t_1}} + \pi_2 e^{-\tau(1-\frac{t_2}{1-t_1})}\right]^{n-n_1} \triangleq [h(\tau)]^{n-n_1} \qquad (56)$$

To obtain the tightest bound we minimize $h(\tau)$, setting the derivative to zero

$$\frac{\partial h}{\partial \tau} = (1-\pi_1-\pi_2)\frac{t_2}{1-t_1}e^{\tau\frac{t_2}{1-t_1}} - \pi_2(1-\frac{t_2}{1-t_1})e^{-\tau(1-\frac{t_2}{1-t_1})} = 0 \ ,$$

we get

$$e^{\tau} = \frac{\pi_2(1-t_1-t_2)}{(1-\pi_1-\pi_2)t_2} = \frac{2+s}{s}$$

Note that $\tau = \ln[(2+s)/s] \geq 0$, as required. Combining these results we obtain

$$h^*(\tau) = \frac{AT}{\Lambda}(t_0+t_2)s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_2}{t_0+t_2}}$$

Turning back to $P_{em}(2)$,

$$\begin{aligned} P_{em}(2) &\leq M_1^\rho M_2^\rho \sum_{n=1}^{\infty} \frac{e^{-\Lambda}\Lambda^n}{n!} \sum_{n_1=0}^{n} \binom{n}{n_1} \pi_1^{n_1} [h^*(\tau)]^{n-n_1} \left[\frac{t_1^{n_1}(1-t_1)^{n-n_1} n^n}{(n-n_1)^{n-n_1} n_1^{n_1}}\right]^\rho \\ &\leq M_1^\rho M_2^\rho \sum_{n=1}^{\infty} \frac{e^{-\Lambda}\Lambda^n}{n!} n \left[(t_1^\rho \pi_1)^{\frac{1}{1+\rho}} + \left((1-t_1)^\rho h^*(\tau)\right)^{\frac{1}{1+\rho}}\right]^{n(1+\rho)} \\ &= B_2 e^{-T(A\hat{E}_2 - \rho R)} \ , \end{aligned}$$

where

$$\begin{aligned} \hat{E}_2 &- \frac{\Lambda}{AT}\left\{1 - \left[(t_1^\rho \pi_1)^{\frac{1}{1+\rho}} + \left((1-t_1)^\rho h^*(\tau)\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\} \\ &- \frac{\Lambda}{AT} - \left[t_1(1+s)^{\frac{1}{1+\rho}} + (t_0+t_2)\left(s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_2}{t_0+t_2}}\right)^{\frac{1}{1+\rho}}\right]^{1+\rho} \end{aligned}$$

To prove that $E_{12} \leq \hat{E}_2$ it remains to show that

$$t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}} - (t_0+t_2)\left(s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_2}{t_0+t_2}}\right)^{\frac{1}{1+\rho}} \geq 0 \ .$$

Let $p = \frac{t_0}{t_0+t_2}$ and $t = (\frac{s}{2+s})^{\frac{1}{1+\rho}}$ then as s increases from 0 to ∞ t increases from 0 to 1. Using these definitions we have to show that

$$r(p,t) = pt + (1-p) - t^p \geq 0 \quad , \quad 0 \leq p \leq 1 \quad , \quad 0 \leq t \leq 1$$

For $0 \leq p \leq 1$ the function $g(t) = t^p$ is concave and $g(1) = 1$, $g'(1) = p$. Hence the graph of $g(t)$ lies below its tangent line at $(1, g(1))$: $1-p+pt$ (see [19, Appendix II]).

Consider the term

$$\begin{aligned} P_{em}(3) &= \sum_{(n_1,n_2,n)\in A_0^c\cap A_1} Q(n_1,n_2,n)\Big[\sum_{m'\neq m} \Gamma_1(n_1,n_2,n)\Big]^\rho \\ &\le M_1^\rho M_2^\rho \sum_{(n_1,n_2,n)} Q(n_1,n_2,n)e^{\tau(nt_1-n_1)}\Gamma_2^\rho(n_1,n_2,n) \end{aligned}$$

For a given n,

$$\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1} Q(n_1,n_2,n)e^{\tau(nt_1-n_1)}\Gamma_2^\rho(n_1,n_2,n) = \frac{e^{-\Lambda}\Lambda^n}{n!}\sum_{n_1=0}^{n}\sum_{n_2=0}^{n-n_1}\binom{n}{n_1}\Big[\pi_1 e^{-\tau(1-t_1)}\Big]^{n_1} h_1(l,n_2)$$

where

$$h_1(l,n_2) \triangleq \binom{l}{n_2} u^{n_2} v^{l-n_2}\left[\frac{l^l p^{n_2}(1-p)^{l-n_2}}{n_2^{n_2}(l-n_2)^{l-n_2}}\right]^\rho ,$$

and $u \triangleq \pi_2 e^{\tau t_1}$, $v \triangleq (1-\pi_1-\pi_2)e^{\tau t_1}$, $p \triangleq t_2/(1-t_1)$, $l \triangleq n-n_1$.

First we maximize $h_1(l,n_2)$ w.r.t. n_2 which yields,

$$\max_{0\le n_2\le l} h_1(l,n_2) = \left[(p^\rho u)^{\frac{1}{1+\rho}} + \Big((1-p)^\rho v\Big)^{\frac{1}{1+\rho}}\right]^{l(1+\rho)} \triangleq r^{l(1+\rho)} .$$

Upper bounding $\sum_{n_2=0}^{n-n_1} h_1(l,n_2)$ by $(n-1)\cdot \max_{0\le n_2\le l} h_1(l,n_2)$ we can now proceed with the outer sum

$$\sum_{n_1=0}^{n}\binom{n}{n_1}\Big[\pi_1 e^{-\tau(1-t_1)}\Big]^{n_1}(r^{1+\rho})^{n-n_1} = \Big(\pi_1 e^{-\tau(1-t_1)} + r^{1+\rho}\Big)^n .$$

To get the tightest bound we optimize the term inside the brackets namely

$$\begin{aligned} g(\tau) &= \pi_1 e^{-\tau(1-t_1)} + \left[(p^\rho\pi_2)^{\frac{1}{1+\rho}} e^{\frac{\tau t_1}{1+\rho}} + \Big((1-p)^\rho(1-\pi_1-\pi_2)\Big)^{\frac{1}{1+\rho}} e^{\frac{\tau t_1}{1+\rho}}\right]^{1+\rho} \\ &= \frac{1}{\Lambda}\left\{t_1(1+s)e^{-\tau(1-t_1)} + \frac{e^{\tau t_1}}{(1-t_1)^\rho}\Big[t_2(2+s)^{\frac{1}{1+\rho}} + t_0 s^{\frac{1}{1+\rho}}\Big]^{1+\rho}\right\} . \end{aligned} \qquad (57)$$

Equating the derivative of $g(\tau)$ to zero and solving for τ we obtain

$$e^\tau = (1+s)\left[\frac{1-t_1}{t_2(2+s)^{\frac{1}{1+\rho}} + t_0 s^{\frac{1}{1+\rho}}}\right]^{1+\rho} . \qquad (58)$$

By Jensen's inequality and the fact that practically we shall consider average energy constraints (for each of the users) in the range $0 \le p_i \le 1/2, i=1,2$ and recalling that $t_0 = (1-p_1)(1-p_2)$, $t_2 = p_1 p_2$ it follows that

$$\left[\frac{t_0}{t_0+t_2}s^{\frac{1}{1+\rho}} + \frac{t_2}{t_0+t_2}(2+s)^{\frac{1}{1+\rho}}\right]^{1+\rho} \le \frac{t_0}{t_0+t_2}s + \frac{t_2}{t_0+t_2}(2+s) \le 1+s,$$

so $\tau = \ln(1+s) / \left[\frac{t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}}}{t_0+t_2}\right]^{1+\rho} \geq 0$ as required.

Combining (58) and (57) we have

$$g^*(\tau) = \frac{AT}{\Lambda}\left[\frac{t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}}}{t_0+t_2}\right]^{(1+\rho)(t_0+t_2)} (1+s)^{t_1} ,$$

- i.e. $P_{em}(3) \leq B_3 e^{-T(A\hat{E}_3 - \rho R)}$, where

$$\hat{E}_3 = \frac{\Lambda}{AT} - \left[\frac{t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}}}{t_0+t_2}\right]^{(1+\rho)(t_0+t_2)} (1+s)^{t_1} \triangleq \frac{\Lambda}{AT} - h_2(t_0, t_2, s) . \tag{59}$$

To prove that $\hat{E}_3 \geq E_{12}$ we need to show that

$$h_2(t_0, t_2, s) \leq [t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}} + t_1(1+s)^{\frac{1}{1+\rho}}]^{1+\rho} \triangleq h_3(t_0, t_2, s) .$$

Let $t = \frac{t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}}}{t_0+t_2} \cdot \frac{1}{(1+s)^{\frac{1}{1+\rho}}}$, so that t increases from $\frac{t_2}{t_0+t_2} 2^{\frac{1}{1+\rho}}$ to 1 as s increases from 0 to ∞.

We have

$$\left[\frac{1}{1+s} h_2(t_0, t_2, s)\right]^{\frac{1}{1+\rho}} =$$

$$\left[\frac{1}{1+s} h_3(t_0, t_2, s)\right]^{\frac{1}{1+\rho}} = \frac{t_0 s^{\frac{1}{1+\rho}} + t_2(2+s)^{\frac{1}{1+\rho}}}{(1+s)^{\frac{1}{1+\rho}}} + t_1 = (1-t_1)t + t_1$$

which completes the proof.

Finally we turn to $P_{em}(4)$,

$$\begin{aligned} P_{em}(4) &= \sum_{(n_1,n_2,n) \in A_0^c \cap A_1^c} Q(n_1, n_2, n) \\ & \sum_{(n_1,n_2,n)} Q(n_1, n_2, n) \chi_0(n_1, n_2, n) \chi_1(n_1, n_2, n) \\ &< \sum_{(n_1,n_2,n)} Q(n_1, n_2, n) e^{\tau_1(nt_1 - n_1)} e^{\tau_2[(n-n_1)\frac{t_2}{1-t_1} - n_2]} \end{aligned}$$

where $\chi_0(n_1, n_2, n)$, $\chi_1(n_1, n_2, n)$ are the indicator functions of the sets A_0^c and A_1^c respectively.

For a given n

$$\sum_{n_1=0}^{n} \sum_{n_2=0}^{n-n_1} Q(n_1, n_2, n) e^{\tau_1(nt_1 - n_1)} e^{\tau_2[(n-n_1)\frac{t_2}{1-t_1} - n_2]} =$$

$$\frac{e^{-\Lambda}\Lambda^n}{n!} \sum_{n_1=0}^{n} \binom{n}{n_1} \left[\pi_1 e^{-\tau_1(1-t_1)}\right]^{n_1} e^{\tau_1 t_1 (n-n_1)}$$

$$\cdot \sum_{n_2=0}^{n-n_1} \binom{n-n_1}{n_2} \left[\pi_2 e^{-\tau_2(1-\frac{t_2}{1-t_1})}\right]^{n_2} \left[(1-\pi_1-\pi_2) e^{\tau_2 \frac{t_2}{1-t_1}}\right]^{n-n_1-n_2}$$

Since the inner sum identifies with $[h(\tau_2)]^{n-n_1}$ in (56) its optimized value is $h^*(\tau_2) = \frac{1}{\Lambda}(t_0 + t_2)s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_0}{t_0+t_2}}$. We consider now the outer sum

$$\sum_{n_1=0}^{n} \binom{n}{n_1} \left[\pi_1 e^{-\tau_1(1-t_1)}\right]^{n_1} \left(h^*(\tau_2)e^{\tau_1 t_1}\right)^{n-n_1} = \left[\pi_1 e^{-\tau_1(1-t_1)} + h^*(\tau_2)e^{\tau_1 t_1}\right]^n \triangleq v^n \tag{60}$$

Setting the derivative of v w.r.t. τ_1 to zero yields

$$e^{\tau_1} = \frac{\pi_1(1-t_1)}{h^*(\tau_2)t_1} = \frac{1+s}{s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_2}{t_0+t_2}}}\,. \tag{61}$$

By [30, Appendix 3A inequality b]

$$s^{\frac{t_0}{t_0+t_2}}(2+s)^{\frac{t_2}{t_0+t_2}} \leq \frac{t_0 s + t_2(2+s)}{t_0+t_2} < (1+s)$$

hence $\tau_1 > 0$ as required. Substituting (61) into (60) we obtain $v = \frac{AT}{\Lambda}s^{t_0}(1+s)^{t_1}(2+s)^{t_2}$ so $P_{em}(4) \leq e^{-\hat{E}_4 T}$ where $\hat{E}_4 = \Lambda/T - As^{t_0}(1+s)^{t_1}(2+s)^{t_2}$.

To show that $\hat{E}_4$ is not less than $E_{12}(\rho, p_1, p_2, q) - \rho R$ it suffices to show that $\hat{E}_4 \geq E_{12}(1, p_1, p_2, q)$. Let $t = s/(1+s)$ and consider the pair of functions f and g defined by

$$\begin{aligned} f &\triangleq [t_0 t^{\frac{1}{2}} + t_2(2-t)^{\frac{1}{2}} + t_1]^2 \\ g &\triangleq t^{t_0}(2-t)^{t_2} \end{aligned}$$

We have

$$\begin{aligned} f'(t) &= [t_0 t^{\frac{1}{2}} + t_2(2-t)^{\frac{1}{2}} + t_1][t_0 t^{-\frac{1}{2}} - t_2(2-t)^{-\frac{1}{2}}] \\ g'(t) &= t_0 t^{t_0-1}(2-t)^{t_2} - t_2 t^{t_0}(2-t)^{t_2-1} \\ f''(t) &= \frac{1}{2}[t_0 t^{-\frac{1}{2}} - t_2(2-t)^{-\frac{1}{2}}]^2 - \frac{1}{2}[t_0 t^{\frac{1}{2}} + t_2(2-t)^{\frac{1}{2}} + t_1][t_0 t^{-\frac{3}{2}} + t_2(2-t)^{-\frac{3}{2}}] \\ g''(t) &= t_0(t_0-1)t^{t_0-2}(2-t)^{t_2} - 2t_0 t_2 t^{t_0-1}(2-t)^{t_2-1} + t_2(t_2-1)t^{t_0}(2-t)^{t_2-2} \end{aligned}$$

Now, $f(1) = g(1) = 1$, $f'(1) = g'(1) = (t_0 - t_2)$ while $g''(1) = (t_0-t_2)^2 - (t_0+t_2) = 2f''(1) < 0$. Next $f(0) = (\sqrt{2}t_2 + t_1)^2$, $g(0) = 0$ - i.e. $f(0) > g(0)$ and as $t \to 0$ $g'(t) = O\left(t^{-(1-t_0)}\right)$ while $f'(t) = O\left(t^{-\frac{1}{2}}\right)$ hence $f'(0) > g'(0)$ since $1/2 \leq t_0 \leq 1$ for $0 \leq p_i \leq 1/2, i = 1,2$. Furthermore, $g''(t) < 0, f''(t) < 0, 0 \leq t \leq 1$ whence both functions are concave. These facts imply that the graph of $f(t)$ lies above that of $g(t)$ and since $\frac{\hat{E}_4 - E_{12}(1,p_1,p_2,q)}{A(1+s)} = f(t) - g(t)$ the proof is complete. ∘

References

[1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete,* vol. 44, pp. 159-175,1978.

[2] R. Ahlswede and N. Cai, "Arbitrarily varying multiple-access channels part I – Ericson's symmetrizability is adequate, Gubner's conjecture is true," *IEEE Trans. Inform. Theory,* vol. 45, No. 2, pp. 742-749, March 1999.

[3] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited: positivity, constraints," *IEEE Trans. Inform. Theory,* vol. IT-34, No. 2, pp. 181-193, March 1988.

[4] I. Csiszár and P. Narayan, "Capacity and decoding rules for the class of arbitrarily varying channels," *IEEE Trans. Inform. Theory,* vol. IT-35, No. 4, pp. 752-769, July 1989.

[5] I. Csiszár and P. Narayan, "Capacity of the Gaussian arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. IT-37, No. 1, pp. 18-26, Jan. 1991.

[6] I. Csiszár, "Arbitrarily varying channels with general alphabets and states," *IEEE Trans. Inform. Theory,* vol. IT-38, No. 6, pp. 1725-1742, November 1992.

[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.

[8] D. Blackwell, L. Breiman and A. J. Thomasian, "The capacity of certain channel classes under random coding," *Ann. Math. Statist.,* vol. 31, pp. 558-567, 1960.

[9] J. A. Gubner, "On the deterministic-code capacity of the multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. IT-36, No. 2, pp. 262-275, March 1990.

[10] J. A. Gubner, "State constraints for the multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. IT-37, No. 1, pp. 27-35, Jan. 1991.

[11] J. A. Gubner and B. L. Hughes, "Nonconvexity of the capacity region of the multiple-access arbitrarily varying channel subject to constraints," *IEEE Trans. Inform. Theory,* vol. IT-41, No. 1, pp. 3-13, Jan. 1995.

[12] J. A. Gubner, "On the capacity region of the discrete additive multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. IT-38, No. 4, pp. 1344-1347, July 1992.

[13] J. -H. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Trans. Inform. Theory,* vol. IT-27, No. 2, pp. 212-226, March 1981.

[14] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inform. Theory,* vol. IT-44, No. 6, pp. 2148-2177, October 1998.

[15] A. Lapidoth and S. Shamai (Shitz), "The Poisson multiple-access channel," *IEEE Trans. Inform. Theory,* vol. IT-44, No. 2, pp. 488-502, March 1998.

[16] Y. M. Kabanov, "The capacity of a channel of the Poisson type," *Thery of Prob. and Its Appl.,* vol. 23, No. 1, pp. 143-147, 1978.

[17] M. H. A. Davis, "Capacity and cut-off rate for Poisson-type channels," *IEEE Trans. Inform. Theory,* vol. IT-26, No. 6, pp. 710-715, November 1980.

[18] M. R. Frey, "Information capacity of the Poisson channel," *IEEE Trans. Inform. Theory,* vol. IT-37, No. 2, pp. 244-256, March 1991.

[19] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel – part I," *IEEE Trans. Inform. Theory,* vol. IT-34, No. 6, pp. 1449-1461, November 1988.

[20] –, "Capacity and error exponent for the direct detection photon channel – part II," *IEEE Trans. Inform. Theory,* vol. IT-34, No. 6 pp. 1462-1471, November 1988.

[21] S. I. Bross, M. V. Burnashev and S. Shamai (Shitz), "Error exponents for the two-user Poisson multiple-access channel," submitted to *IEEE Trans. Inform. Theory.*

[22] M. V. Burnashev and Y. A. Kutoyants, "On sphere-packing bound, capacity and related results for the Poisson channel," *Probl. Inform. Transm.,* vol. 35, No. 2, pp. 3-22 (1999).

[23] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory,* vol. IT-36, No. 5, pp. 1019-1030, Sept. 1990.

[24] R. G. Gallager, "Energy limited channels: Coding, multiaccess, and spread spectrum," Laboratory for Information and Decision Systems, Mass. Inst. Technol. Tech. Rep. LIDS-P-1714, Nov. 1988; see also in *Proc. 1988 Conf. Information Science and Systems* (Princeton, NJ, Mar. 1988), p. 372.

[25] –, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[26] P. Narayan and D. L. Snyder, "Signal set design for band-limited memoryless multiple access channels with soft decision demodulation," *IEEE Trans. Inform. Theory,* vol. IT-33, No. 4, pp. 539-556, July 1987.

[27] S. Verdú, "Multiple-access channels with point-process observations: optimum demodulation," *IEEE Trans. Inform. Theory,* vol. IT-32, No. 5, pp. 642-651, Sept. 1986.

[28] Y. A. Kutoyants, *Parameter Estimation for Stochastic Processes.* Berlin: Heldermann, 1984.

[29] D. L. Snyder, *Random Point Processes.* New York: Wiley, 1975.

[30] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hill 1979.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[32] A. W. Roberts and D. E. Varberg, *Convex Functions.* New York: Academic Press, 1973.

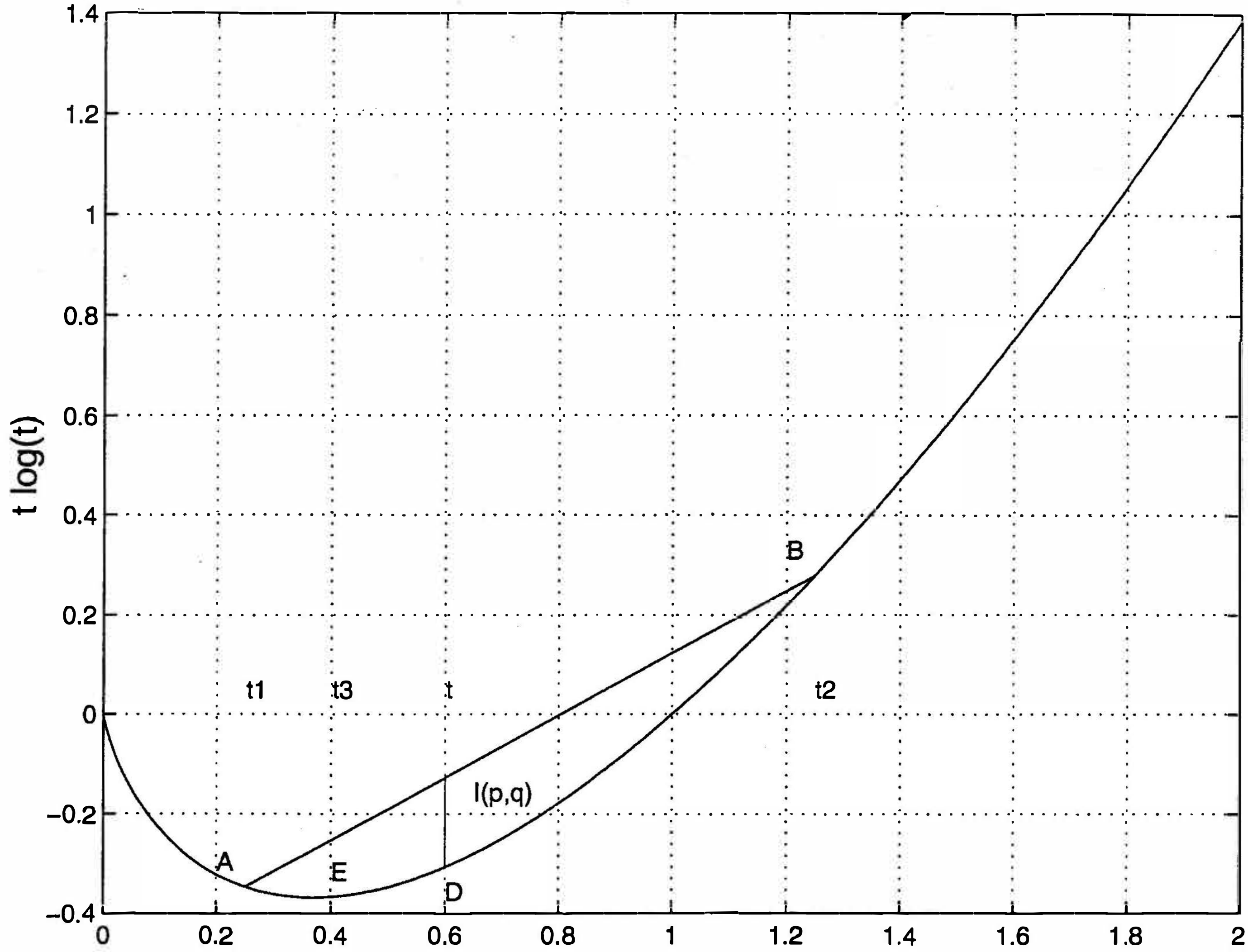

Figure 1: Geometrical interpretation of $I(p, \Lambda)$. $t_1 = s_0 + q$, $t_2 = s_0 + q + 1$, $t = s_0 + q + p$

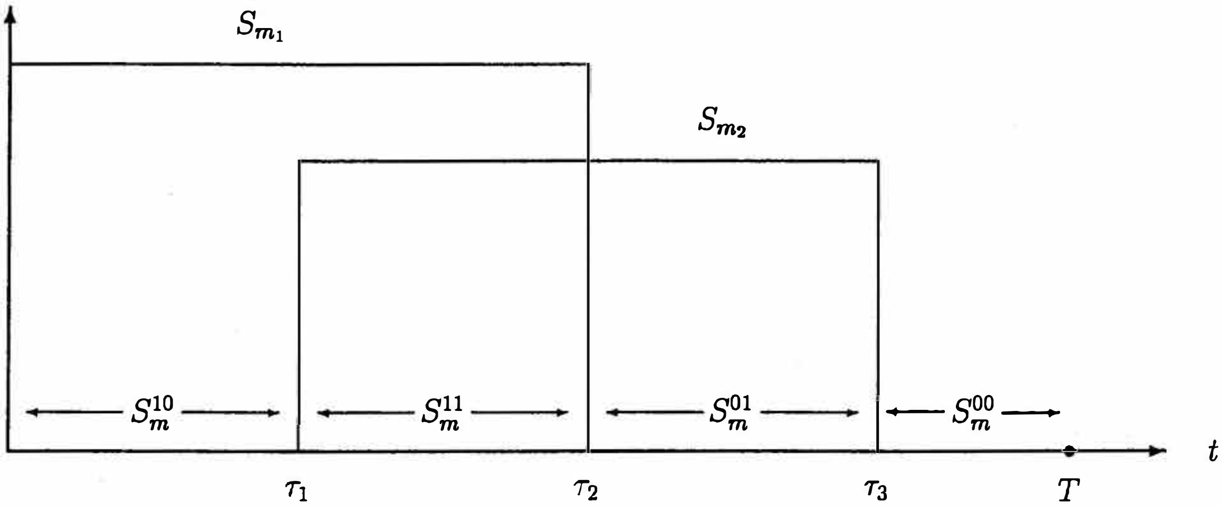

Figure 2: Schematic diagram of a two-user code waveforms.

The *Center for Communication and Information Technologies* (CCIT)

is managed by the Department of Electrical Engineering.

This Technical Report is listed also as

EE PUB #1260, September 2000.