A Pairwise Error Probability Bound for the Exponential-Server Timing Channel

Aaron B. Wagner and Venkat Anantharam
Dept. of Electrical Engineering and Computer Sciences
University of California
Berkeley, CA 94720 USA
{awagner,ananth}@eecs.berkeley.edu

***Abstract*— We exhibit upper and lower bounds on the pairwise error probability of the exponential-server timing channel in terms of an appropriately-defined distance between codewords. We show that this distance plays a crucial role in determining the reliability function at low rates. In particular, by lower bounding the minimum distance of good low-rate codes, we provide an improved lower bound on the reliability function of the channel at rate zero. This improved bound proves that at low rates, the exponential-server timing channel is strictly more reliable than is the related Poisson channel with zero dark current, answering an open question posed by Arikan. Some remarks are also made about using the results of this paper to prove an improved upper bound on the reliability function at rate zero.**

I. Introduction

We consider the exponential-server timing channel (ESTC), as introduced by Anantharam and Verdú [1], in which the sender chooses the arrival times of identical jobs to a $\cdot/M/1$ queue, while the receiver observes their departure times. Timing channels such as the ESTC have been studied in the context of covert communication [2]; by modulating the times at which it performs routine tasks, the sender can send information to the receiver so that it is unclear to a casual observer that communication has occurred. We do not consider this intriguing application here but rather focus on the channel model itself.

We shall assume that the service discipline is first-in-first-out, and that the queue is initially empty. Neither of these assumptions is crucial; see Anantharam and Verdú [1] for a discussion of these assumptions and for a discussion of the cost of transmission for the channel. Unlike conventional channels, the transmission and reception times of the ESTC can be quite different, and the reception time depends on the realization of the channel. This results in multiple reasonable definitions of a block code. We shall resolve this ambiguity here by considering only the *window codes* of Sundaresan and Verdú [3].

A (n, M, T) (window) code is a collection of M codewords, each of which is a vector of n nonnegative interarrival times $(v_1, \ldots, v_n)$ such that the kth arrival occurs at time $\sum_{i=1}^{k} v_i$, and the nth arrival occurs before time T; and a decoder that selects a codeword or declares an error for each possible channel output. The (data) rate of a (n, M, T) code is defined to be $(1/T) \log M$. The arrival rate is n/T. We exclusively use logarithms with base e.

The capacity achievable using window codes is shown to be μ/e nats/time by Sundaresan and Verdú [3], where $1/\mu$ is the mean service time. Anantharam and Verdú [1] find the same capacity using a different block code definition. Arikan [4] proves that the cutoff rate is $\mu/4$, and gives upper and lower bounds on the reliability function that coincide at rates above $(\mu/4)\log 2$. These bounds are reproduced in Section III.

Arikan points out that these bounds provide for an interesting comparison between the ESTC and the Poisson channel with zero dark current and a peak power constraint of μ. The latter models a direct-detection optical channel; the input is a nonnegative waveform $\lambda(t)$, which is upper bounded by μ, while the output is a Poisson process with intensity $\lambda(t)$. The capacity of this channel is also μ/e nats/time, and its reliability function is completely known [5], [6] and coincides with Arikan's lower bound on the reliability function of the ESTC. Thus the ESTC is at least as reliable as is the Poisson channel, and their reliability functions coincide at rates between $(\mu/4)\log 2$ and their common capacity. Arikan [4] posed the problem of determining whether the ESTC is strictly more reliable at low rates than is the Poisson channel.

It is instructive to begin a study of this problem by considering the binary symmetric and Gaussian channels. These channels have a natural distance metric over the channel inputs (Hamming distance and Euclidean distance, respectively), such that the maximum-likelihood (ML) error probability between a pair of codewords only depends on the distance between them. For low-rate codes over these channels, the error probability of a codeword is well-approximated by the probability that it is confused with its "nearest neighbor," measured according to this metric. The reliability function at low rates in turn is governed by the greatest distance by which the codewords can be separated in a sequence of codes with positive asymptotic rate.

We show that much of this reasoning applies to the ESTC. Unfortunately, the standard estimates of the pairwise error probability under ML detection are intractable for this channel. To bypass this, we introduce a simple suboptimal decoder for which the pairwise error probability can be readily estimated. We show that at low arrival rates, this decoder performs well, and the pairwise error probability is governed by the distance between the codewords, measured according to a particular metric. These results are contained in Section II.

In Section III we show that the problem of determining the ESTC's reliability function at low rates can be reduced to a combinatorial sphere-packing problem involving our distance metric. By lower bounding the minimum distance of good low-rate codes, we prove that the ESTC is indeed strictly more reliable at low rates than is the Poisson channel. Some concluding remarks are made in Section IV.

II. Pairwise Error Probability Bounds

The exponentially-distributed service times make the ML decoder for the ESTC particularly simple. Given the received departure times, the ML decoder first excludes any codeword for which a departure occurs before the corresponding arrival; we call such a codeword *infeasible* for this output. For each feasible codeword, the decoder computes the service times that would cause the observed departure times. It then selects the codeword whose service times have the smallest sum.

Unfortunately, the simplicity of this decoder does not translate into ease of analysis. To proceed, we introduce a suboptimal decoder that approximates ML detection. Observe that every pair of codewords $u^n \neq v^n$ partitions the output space into four sets: the outputs for which u^n is feasible but v^n is not, the outputs for which v^n is feasible but u^n is not, the outputs for which neither are feasible, and the outputs for which they both are. In general, the ML decoder will split this last set between the decision regions of u^n and v^n; our suboptimal decoder will place the entire set in one of the two. To specify which, we use the following definition.

Definition 1: To each codeword $(u_1, \ldots, u_n)$ we associate the arrival process $u(t)$ defined by

$$u(t) := \sup\left\{0 \leq i \leq n : \sum_{j=1}^{i} u_j \leq t\right\},$$

where an empty sum is defined to be zero. For two codewords u^n and v^n, we say that u^n *leads* v^n by

$$\mathcal{L}(u^n, v^n) := \int_0^\infty 1(u(t) > v(t))\, dt.$$

Then our *pseudo-ML decoder* for the codeword pair u^n, v^n operates as follows: if $\mathcal{L}(u^n, v^n) > \mathcal{L}(v^n, u^n)$, then the decoding region for v^n consists of all outputs for which v^n is feasible, and all others are assigned to u^n. Similarly, if $\mathcal{L}(v^n, u^n) > \mathcal{L}(u^n, v^n)$, then the decoding region for u^n consists of all outputs for which u^n is feasible, and all others are assigned to v^n. If $\mathcal{L}(u^n, v^n) = \mathcal{L}(v^n, u^n)$, then we arbitrarily assign the outputs for which both codewords are feasible to one of the two codewords. Although the choice might affect the error probability of the decoder, it will not affect our bound on it.

To evaluate this bound, we use the following lemma, whose proof follows immediately from Chernoff's bound.

Lemma 1: If $\{S_i\}_{i=1}^\infty$ are i.i.d. exponential with mean $1/\mu$, then for all $a > 1/\mu$, all $0 < b < 1/\mu$, and all $n \geq 1$,

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} S_i \geq a\right) \leq \exp(-n(a\mu - 1 - \log(a\mu))),$$

and

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} S_i \leq b\right) \leq \exp(-n(b\mu - 1 - \log(b\mu))).$$

More useful than the $\mathcal{L}$ quantity defined above is the following symmetrized version.

Definition 2: For two codewords u^n and v^n, let

$$\overline{\mathcal{L}}(u^n, v^n) := \max(\mathcal{L}(u^n, v^n), \mathcal{L}(v^n, u^n)).$$

For a code $\mathcal{C}$ consisting of codewords $(u_1^n, \ldots, u_M^n)$, let

$$\overline{\mathcal{L}}(\mathcal{C}) := \inf_{i \neq j} \overline{\mathcal{L}}(u_i^n, u_j^n).$$

Theorem 1: Let u^n and v^n be the two codewords in a $(n, 2, T)$ code that uses a pseudo-ML decoder, and let P_u and P_v be their probabilities of error. If $\overline{\mathcal{L}}(u^n, v^n) > n/\mu$, where $1/\mu$ is the mean service time, then

$$\max(P_u, P_v) \leq \exp(-n(\mu\overline{\mathcal{L}}(u^n, v^n)/n - 1 - \log(\overline{\mathcal{L}}(u^n, v^n)\mu/n))).$$

Proof: When using a pseudo-ML decoder, $P_u \cdot P_v = 0$, so suppose without loss of generality that $P_v = 0$. Then $\mathcal{L}(u^n, v^n) = \overline{\mathcal{L}}(u^n, v^n)$, and given that u^n was transmitted, an error occurs if and only if v^n is feasible for the resulting output. But this requires that the server be busy whenever $u(t) > v(t)$. So if $S_1, \ldots, S_n$ are the service times,

$$P_u \leq \Pr\left(\sum_{i=1}^{n} S_i \geq \mathcal{L}(u^n, v^n)\right).$$

The result then follows from Lemma 1. □

A corollary of the theorem is that for the ML decoder, assuming that $\overline{\mathcal{L}}(u^n, v^n) > n/\mu$,

$$\frac{P_u + P_v}{2} \leq \exp\left(-n\left(\mu\overline{\mathcal{L}}(u^n, v^n)/n - 1 - \log\left(\overline{\mathcal{L}}(u^n, v^n)\mu/n\right)\right)\right).$$

The $\overline{\mathcal{L}}$ quantity is the distance mentioned in the introduction. Two of its properties are listed in the next result, whose straightforward proof is omitted.

Lemma 2:

(a) $\overline{\mathcal{L}}$ is a metric on the space of codewords with n points, for each n.

(b) If u^n and v^n are codewords and w^n is defined by $w(t) = \min(u(t), v(t))$, then

$$\overline{\mathcal{L}}(u^n, w^n) = \mathcal{L}(u^n, v^n),$$
$$\overline{\mathcal{L}}(v^n, w^n) = \mathcal{L}(v^n, u^n).$$

Observe that if u^n and v^n are two codewords and w^n is as in (b), then w^n, when viewed as a channel output, is feasible for both u^n and v^n. A ML detector will assign it to u^n if $\mathcal{L}(u^n, w^n) < \mathcal{L}(v^n, w^n)$, and to v^n if $\mathcal{L}(v^n, w^n) < \mathcal{L}(u^n, w^n)$, since $\mathcal{L}(x^n, y^n)$ is the sum of the service times required to cause output y^n when the input is x^n. Since $\mathcal{L}(u^n, w^n) = \mathcal{L}(u^n, v^n)$ and $\mathcal{L}(v^n, w^n) = \mathcal{L}(v^n, u^n)$, a pseudo-ML decoder treats all outputs that are feasible for both codewords identically to w^n. We expect the resulting loss to be small since, regardless of which codeword is sent, we expect w^n to be the dominating point of the set of outputs for which both codewords are feasible. The next theorem bounds this loss.

Theorem 2: Let u^n and v^n be two codewords. For any decoder, if P_u and P_v are their probabilities of error, then

$$\max(P_u, P_v) \geq \frac{1}{2}\exp\left(-\mu\overline{\mathcal{L}}(u^n, v^n)\right),$$

where $1/\mu$ is the mean service time.

Proof: Let $D_u \subset \mathbb{R}_+^n$ be the decision region for u^n and let $D_v \subset \mathbb{R}_+^n$ be the decision region for v^n. Define the codeword w^n as in Lemma 2(b), and let

$$Y_u = \{y^n \in \mathbb{R}_+^n : u^n \text{ is feasible for the output } y^n\},$$

and define Y_v and Y_w similarly. Then

$$\begin{aligned}
P_u &= \int_{D_u^c \cap Y_u} \mu^n \exp\left(-\mu\overline{\mathcal{L}}(u^n, y^n)\right)\, dy^n, \\
&\geq \int_{D_u^c \cap Y_w} \mu^n \exp\left(-\mu\left(\overline{\mathcal{L}}(u^n, w^n) + \overline{\mathcal{L}}(w^n, y^n)\right)\right)\, dy^n, \\
&= \exp(-\mu\mathcal{L}(u^n, v^n)) \cdot \\
&\qquad \int_{D_u^c \cap Y_w} \mu^n \exp\left(-\mu\overline{\mathcal{L}}(w^n, y^n)\right)\, dy^n, \\
&\geq \exp(-\mu\overline{\mathcal{L}}(u^n, v^n)) \cdot \\
&\qquad \int_{D_u^c \cap Y_w} \mu^n \exp\left(-\mu\overline{\mathcal{L}}(w^n, y^n)\right)\, dy^n,
\end{aligned}$$

where we have used Lemma 2. Similarly,

$$\begin{aligned}
P_v &\geq \exp(-\mu\overline{\mathcal{L}}(u^n, v^n)) \cdot \\
&\qquad \int_{D_v^c \cap Y_w} \mu^n \exp\left(-\mu\overline{\mathcal{L}}(w^n, y^n)\right)\, dy^n.
\end{aligned}$$

But

$$\int_{Y_w} \mu^n \exp(-\mu\overline{\mathcal{L}}(w^n, y^n))\, dy^n = 1.$$

So since D_u and D_v are disjoint,

$$\begin{aligned}
&\int_{D_u^c \cap Y_w} \mu^n \exp(-\mu\overline{\mathcal{L}}(w^n, y^n))\, dy^n + \\
&\qquad \int_{D_v^c \cap Y_w} \mu^n \exp(-\mu\overline{\mathcal{L}}(w^n, y^n))\, dy^n \geq 1.
\end{aligned}$$

Therefore at least one of these integrals must be at least 1/2. □

III. Low-Rate Reliability

We illustrate the efficacy of the bounds in the previous section by using them to establish an improved lower bound on the reliability function of the ESTC at low rates.

Let $P_e(R, T)$ be the infimum of the maximum probability of error over all codes that are (n, M, T) for some $n \geq 1$ and $M \geq e^{RT}$. Then the *reliability function* (or *error exponent*) of the channel at rate $R > 0$ is defined by

$$E(R) := \limsup_{T\to\infty} -\frac{1}{T}\log P_e(R, T).$$

The *zero-rate error exponent* is defined by

$$E(0) := \sup_{R>0} E(R).$$

By modifying the random-coding and sphere-packing bounds [7] to handle the unique features of this channel, Arikan [4] proves the following.

Proposition 1 ([4]): Let $E_{sp}(R)$ be defined parametrically for $0 < R < \mu/e$ by

$$\begin{aligned}
E_{sp}(R) &= \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}}[\rho - \log(1+\rho)], \\
R &= \frac{\mu}{(1+\rho)^{(1+\rho)/\rho}}\log(1+\rho)^{1/\rho},
\end{aligned}$$

where ρ ranges over $(0, \infty)$, and let $E_{sp}(0) = \mu$. For $R_c := (\mu/4)\log 2 < R < \mu/e$, let $E_r(R) = E_{sp}(R)$, while for $0 \leq R \leq R_c$, let $E_r(R) = \mu/4 - R$. Then the reliability function of the exponential-server timing channel with mean service time $1/\mu$ satisfies

$$E_r(R) \leq E(R) \leq E_{sp}(R)$$

for all $0 \leq R < \mu/e$.

Since $E_r(R)$ and $E_{sp}(R)$ coincide for R in $[R_c, \mu/e]$, the reliability function is determined at these rates. The reliability function is currently not known at rates below R_c, and in particular $E(0)$ is not known: one only has the bounds $\mu/4 \leq E(0) \leq \mu$. But determining $E(0)$ might reveal whether the ESTC is more reliable at low rates than is the Poisson channel, whose reliability function is known to be exactly $E_r(R)$ [5], [6]. One expects that it is, since the random-coding exponent is rarely tight at low rates. In particular, one can show using Gallager's expurgated exponent [7] that the random-coding bound lies strictly below the reliability function at sufficiently low rates for any discrete memoryless channel with nonzero capacity. Unfortunately, one cannot evaluate the standard expurgated bound for the ESTC. We shall instead use the results from the last section to lower bound $E(0)$.

Definition 3: For $n \in \mathbb{N}$ and $x \in (0, 1]$, let

$$M(n, x) = \sup\left\{M : \text{ there exists a } (n, M, 1) \text{ code } \mathcal{C} \text{ with } \overline{\mathcal{L}}(\mathcal{C}) \geq x\right\}.$$

It is not difficult to verify that $M(n, 1) = n + 1$ for all n and $M(n, x) < \infty$ for all n and x. Define also

$$L_0 = \sup\left\{x \in [0, 1] : \limsup_{n\to\infty}\frac{1}{n}\log M(n, x) > 0\right\}.$$

We call L_0 the *maximum separable distance* of the channel. We shall show later that it satisfies $L_0 \geq 1/2$. First we link it to the zero-rate error exponent.

Proposition 2: The zero-rate error exponent of the exponential-server timing channel with mean service time $1/\mu$ satisfies $E(0) = \mu L_0$.

Proof: We shall first show that $E(0) \geq \mu L_0$. This is obvious if $L_0 = 0$. Otherwise, let $0 < \epsilon < L_0$. By the definition of L_0, there exists a sequence of codes $\{\mathcal{C}_n\}$ such that $\mathcal{C}_n$ is $(n, M_n, 1)$ for each n, $\overline{\mathcal{L}}(\mathcal{C}_n) \geq L_0 - \epsilon$ for all n, and

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n = R > 0.$$

Let $0 < \lambda < \mu$, and for each n, redefine $\mathcal{C}_n$ to be the $(n, M_n, n/\lambda)$ code obtained by dilating time in the original codewords by a factor of n/λ. Consider decoding $\mathcal{C}_n$ using a pseudo-ML decoder: we make a binary decision between all distinct pairs of codewords, and if one codeword wins all the comparisons in which it participates, we select it. Otherwise, we declare an error.

By choosing λ sufficiently small, we may assume that $\mu(L_0 - \epsilon)/\lambda > 1$. Then using the union bound, Lemma 1, and the fact that $x - 1 - \log x$ is increasing on $[1, \infty)$, we have that the error probability of any codeword in $\mathcal{C}_n$ is upper bounded by

$$(M_n - 1)\exp(-n(\mu(L_0 - \epsilon)/\lambda - 1 - \log(\mu(L_0 - \epsilon)/\lambda))).$$

If p_n is the maximum error probability of all codewords in $\mathcal{C}_n$, then

$$p_n \leq (M_n - 1)\exp(-n(\mu(L_0 - \epsilon)/\lambda - 1 - \log(\mu(L_0 - \epsilon)/\lambda))).$$

Since the asymptotic rate of $\{\mathcal{C}_n\}$ is positive,

$$\limsup_{n\to\infty} \frac{1}{n/\lambda} \log M_n = \lambda R > 0,$$

we have

$$\begin{aligned} E(0) &\geq \liminf_{n\to\infty} -\frac{1}{n/\lambda} \log p_n, \\ &\geq -\lambda R + \mu(L_0 - \epsilon) - \lambda - \lambda \log(\mu(L_0 - \epsilon)/\lambda). \end{aligned}$$

Taking $\lambda \to 0$ and $\epsilon \to 0$ establishes that $E(0) \geq \mu L_0$. Now let $0 < \delta < E(0)$ and let $\{\mathcal{C}_k\}$ be a sequence of codes such that $\mathcal{C}_k$ is (n_k, M_k, T_k) with maximum error probability p_k satisfying

$$T_k \uparrow \infty, \tag{a}$$

$$\lim_{k\to\infty} -\frac{1}{T_k} \log p_k \geq E(0) - \delta, \tag{b}$$

and

$$\lim_{k\to\infty} \frac{1}{T_k} \log M_k > 0. \tag{c}$$

If $\liminf T_k/n_k = 0$, then we can remedy this by choosing $K > 1$ such that $1 + K \log K - K > (E(0) - \delta)/\mu$, and discarding any points beyond the first $\lceil K\mu T_k \rceil =: N_k$ in each codeword of $\mathcal{C}_k$. If p'_k is the maximum error probability of the new code then by Lemma 1,

$$\begin{aligned} p'_k &\leq p_k + \Pr\left(\frac{1}{N_k}\sum_{i=1}^{N_k} S_i \leq \frac{T_k}{N_k}\right), \\ &\leq p_k + \exp(-N_k(1/K - 1 - \log(1/K))), \\ &\leq p_k + \exp(-T_k(E(0) - \delta)), \end{aligned}$$

so that (b) above holds with p'_k in place of p_k. Therefore we shall assume that $\liminf T_k/n_k > 0$.

For each k there exists a pair of codewords u^{n_k} and v^{n_k} in $\mathcal{C}_k$ such that $\overline{\mathcal{L}}(u^{n_k}, v^{n_k}) = \overline{\mathcal{L}}(\mathcal{C}_k)$. Then by Theorem 2,

$$p_k \geq \max(P_u, P_v) \geq \frac{1}{2}\exp(-\mu\overline{\mathcal{L}}(\mathcal{C}_k)).$$

Thus

$$0 < E(0) - \delta \leq \mu \cdot \liminf_{k\to\infty} \frac{\overline{\mathcal{L}}(\mathcal{C}_k)}{T_k}.$$

Since $M_k \to \infty$ by (a) and (c), and $\liminf_{k\to\infty} \overline{\mathcal{L}}(\mathcal{C}_k)/T_k > 0$, it follows that $n_k \to \infty$. This combined with the observation

$$\liminf_{k\to\infty} \frac{\log M_k}{n_k} \geq \liminf_{k\to\infty} \frac{T_k}{n_k} \cdot \liminf_{k\to\infty} \frac{1}{T_k} \log M_k > 0$$

implies

$$\liminf_{k\to\infty} \frac{\overline{\mathcal{L}}(\mathcal{C}_k)}{T_k} \leq L_0,$$

Thus $E(0) - \delta \leq \mu L_0$, but $\delta > 0$ was arbitrary. □

The problem of determining $E(0)$ for the ESTC has been reduced to the combinatorial problem of determining L_0. Next we shall show that $L_0 \geq 1/2$ using a random coding argument. First we require the following lemma about random walks.

Lemma 3: Let $\{X_t\}_{t=0}^{\infty}$ be a continuous-time random walk on $\mathbb{Z}$ with $X_0 = 0$ and transition rates

$$q(n, n+1) = q(n, n-1) = \lambda > 0 \text{ for all } n \in \mathbb{Z}.$$

Let $\tau_t = \int_0^t 1(X_s = 0)\, ds$. Then for all $\delta > 0$, there exists $C, \gamma > 0$ such that for all $t > 0$,

$$\Pr(\tau_t/t > \delta) \leq C\exp(-\gamma t).$$

Proof: Let $U_0 = 0$ and for $i \geq 1$, let

$$\begin{aligned} V_i &= \inf\{t > U_{i-1} : X_t \neq 0\}, \\ U_i &= \inf\{t > V_i : X_t = 0\}, \end{aligned}$$

and for $i \geq 1$, let

$$\begin{aligned} Y_i &= V_i - U_{i-1}, \\ Z_i &= U_i - V_i. \end{aligned}$$

Let $N_t = \sup\{n : U_n \leq t\}$. Then

$$\Pr(\tau_t/t > \delta) \leq \Pr(N_t > \delta\lambda t/2) + \Pr(\tau_t/t > \delta, N_t \leq \delta\lambda t/2).$$

Now $E[Z_1] = \infty$, so for any $M \in \mathbb{N}$, there exists M' such that if $Z'_i = \min(Z_i, M')$ then $M < E[Z'_1] < \infty$, and there exists $\epsilon > 0$ such that

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} Z_i \leq M\right) \leq \Pr\left(\frac{1}{n}\sum_{i=1}^{n} Z'_i \leq M\right) \leq \exp(-\epsilon n).$$

by Hoeffding's inequality [8]. Write $T = \lceil \delta\lambda t/2 \rceil$. Then there exists $\epsilon > 0$ such that

$$\Pr(N_t > \delta\lambda t/2) \leq \Pr\left(\sum_{i=1}^{T} Z_i \leq t\right) \leq \exp(-\epsilon\delta\lambda t/2).$$

And by Lemma 1, there exists $\alpha > 0$ such that for all sufficiently large t,

$$\Pr(\tau_t/t > \delta, N_t \le \delta\lambda t/2) \le \Pr\left(\sum_{i=1}^{T+1} Y_i > t\delta\right) \le \exp(-\alpha\delta\lambda t/2).$$

□

Define a *n-point Poisson process with rate* λ to be a random vector U^n such that $\{U_i\}_{i=1}^n$ are i.i.d. exponential with mean $1/\lambda$. Define a *constrained n-point Poisson process with rate* λ to be a random vector V^n with distribution

$$\Pr(V^n \in A) = \Pr\left(U^n \in A | n/\lambda - \sqrt{n} < \sum_{i=1}^n U_i < n/\lambda + \sqrt{n}\right),$$

where U^n is a (unconstrained) n-point Poisson process with rate λ.

Proposition 3: The maximum separable distance satisfies $L_0 \ge 1/2$.

Proof: Let U^n and V^n be independent n-point Poisson processes with rate λ, and define the event

$$A_n = \left\{\sum_{i=1}^n U_i \ge n/\lambda - \sqrt{n}, \sum_{i=1}^n V_i \ge n/\lambda - \sqrt{n}\right\}.$$

Write $T_n = n/\lambda + \sqrt{n}$. Then if $\tilde{U}^n$ and $\tilde{V}^n$ are independent constrained n-point Poisson processes with rate λ,

$$\Pr\left(\overline{\mathcal{L}}(\tilde{U}^n, \tilde{V}^n) < (1/2 - \delta)T_n\right) \le \frac{\Pr\left(\overline{\mathcal{L}}(U^n, V^n) < (1/2 - \delta)T_n, A_n\right)}{\Pr(n/\lambda - \sqrt{n} < \sum_{i=1}^n U_i < n/\lambda + \sqrt{n})^2}.$$

Let X_t and τ_t be as in the previous lemma, coupled so that U^n and V^n are the interevent times of the first n up-movements and the first n down-movements of X_t, respectively. Then

$$\Pr\left(\overline{\mathcal{L}}(U^n, V^n) < (1/2 - \delta)T_n, A_n\right) \le \Pr\left(\tau_{n/\lambda - \sqrt{n}} > n/\lambda - \sqrt{n} - (1 - 2\delta)(n/\lambda + \sqrt{n})\right).$$

By the previous lemma, there exists $C, \gamma > 0$ so that

$$\Pr\left(\tau_{n/\lambda - \sqrt{n}} > n/\lambda - \sqrt{n} - (1 - 2\delta)(n/\lambda + \sqrt{n})\right) \le C\exp(-\gamma n).$$

This combined with the observation that

$$\inf_n \Pr\left(n/\lambda - \sqrt{n} < \sum_{i=1}^n U_i < n/\lambda + \sqrt{n}\right) > 0$$

shows that

$$\Pr\left(\overline{\mathcal{L}}(\tilde{U}^n, \tilde{V}^n) < (1/2 - \delta)T_n\right) \le C\exp(-\gamma n)$$

for all n. Then for all sufficiently large n, a $(n, \exp(\gamma n/4), T_n)$ code $\mathcal{C}_n$ consisting of codewords chosen independently and identically distributed to $\tilde{U}^n$ will satisfy

$$\overline{\mathcal{L}}(\mathcal{C}_n) \ge (1/2 - \delta)T_n$$

with high probability. Normalizing the code to have block-length 1 shows that $L_0 \ge 1/2 - \delta$. Since δ was arbitrary, the proof is complete. □

Theorem 3 follows immediately from Proposition 2 and Proposition 3.

Theorem 3: The zero-rate error exponent of the exponential-server timing channel with mean service time $1/\mu$ satisfies $E(0) \ge \mu/2$.

IV. Concluding Remarks

It would follow that $L_0 = 1/2$ and hence that $E(0) = \mu/2$ if one could show that for all $\delta > 0$, there exists a function $f_\delta(n)$ that grows subexponentially with n such that any $(n, M, 1)$ code $\mathcal{C}$ with $M \ge f_\delta(n)$ satisfies $\overline{\mathcal{L}}(\mathcal{C}) < 1/2 + \delta$. We are unable to prove or disprove this assertion, but we note that a code described by Arikan [4] shows that $f_\delta(n)$ must grow at least linearly in order for the assertion to hold: consider the $(n, n+1, 1)$ code $\mathcal{C}_n$ such that the mth codeword places $m-1$ jobs at time 0 and $n+1-m$ jobs at time 1, for $1 \le m \le n+1$. This code satisfies $\overline{\mathcal{L}}(\mathcal{C}_n) = 1$ for each n and number of codewords grows linearly with the number of points.

Acknowledgment

This research was supported by DARPA Grants F30602-00-2-0538 and N66001-00-C-8062, by Grant N00014-1-0637 from the Office of Naval Research, and by Grant SBR-9873086 and a Graduate Research Fellowship from the National Science Foundation.

References

[1] V. Anantharam and S. Verdú, “Bits through queues,” *IEEE Trans. Inform. Theory*, vol. 42, no. 1, pp. 4–18, Jan. 1996.
[2] J. Giles and B. Hajek, “An information-theoretic and game-theoretic study of timing channels,” *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2455–2477, Sept. 2002.
[3] R. Sundaresan and S. Verdú, “Robust decoding for timing channels,” *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 405–419, Mar. 2000.
[4] E. Arikan, “On the reliability exponent of the exponential timing channel,” *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1681–1689, Jun. 2002.
[5] A. D. Wyner, “Capacity and error exponent for the direct detection photon channel—Part I,” *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1449–1461, Nov. 1988.
[6] A. D. Wyner, “Capacity and error exponent for the direct detection photon channel—Part II,” *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1462–1471, Nov. 1988.
[7] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.
[8] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, Oxford University Press, Oxford, 2nd edition, 1992.