Entropy and the Timing Capacity of Discrete Queues

Balaji Prabhakar, *Member, IEEE,* and Robert Gallager, *Life Fellow, IEEE*

***Abstract*—Queueing systems which map Poisson input processes to Poisson output processes have been well-studied in classical queueing theory. This paper considers two discrete-time queues whose analogs in continuous-time possess the Poisson-in–Poisson-out property. It is shown that when packets arriving according to an arbitrary ergodic stationary arrival process are passed through these queueing systems, the corresponding departure process has an entropy rate no less (some times strictly more) than the entropy rate of the arrival process. Some useful by-products are discrete-time versions of: i) a proof of the celebrated Burke's theorem (Burke, 1956), ii) a proof of the uniqueness, amongst renewal inputs, of the Poisson process as a fixed point for exponential server queues (Anantharam, 1993), and iii) connections with the timing capacity of queues (Anantharam and Verdú, 1996).**

***Index Terms*—Entropy, Palm theory, queueing systems, timing capacity.**

I. Introduction

SEVERAL results in classical queueing theory state that certain queueing systems have the Poisson-in–Poisson-out property. That is, if the arrival process to such a queueing system is Poisson, and it is stable (arrival rate $<$ service rate), then the equilibrium departure process from the queueing system is also Poisson. These systems include, for example, the first-come-first-served (FCFS) exponential server queue (symbolically, the $\cdot$/M/1 queue); a queue which dispenses independent and identically distributed (i.i.d.) services with a general distribution and has either of the following service disciplines: 1) last-come-first-served (LCFS) with pre-emptive resume (the $\cdot$/GI/1-LCFS queue), 2) processor sharing (the $\cdot$/GI/1-PS queue); infinite server queues where the service times are i.i.d. and arbitrarily distributed (the $\cdot$/GI/∞ queue); Jackson Networks; and others which incorporate traffic of different classes. Details of these results may be found, for example, in [10], [17].

We shall show that the discrete versions of some of these queueing systems are *entropy increasing* in the following sense: When an arbitrary stationary and ergodic arrival process is passed through such a queueing system, the corresponding equilibrium departure process has an entropy rate no less (and sometimes strictly more) than that of the arrival process.

We explore the connection of entropy-increasing properties with the timing capacity of queues, as considered in the recent paper of Anantharam and Verdú [2]. Anantharam and Verdú consider the problem of a sender transmitting messages encoded in the arrival times of packets to a queue. The receiver tries to decode the message by observing the departure times of the packets, the randomness of the packet service times corrupting the transmitted message. Discrete-time analogs of this model were considered by Bedekar and Azizoğlu [4], and Thomas [16]. We consider the timing capacity of the $\cdot$/GI/1-FCFS queue and rederive some formulas obtained earlier in [4], [16].

Repeated use is made of two basic techniques: one queueing-theoretic and the other information-theoretic. The queueing-theoretic technique consists of comparing the statistical evolution of the queue in forward and in reversed time. This technique has been used to good effect in the study of reversible and quasi-reversible queueing systems, of which the queues considered here are examples (see [10] for a detailed analysis of reversible queueing networks). The information-theoretic technique consists of the following basic fact. Let $\mathcal{S}$ and $\mathcal{R}$ be finite or countable sets, $f\colon \mathcal{S} \to \mathcal{R}$ be a bijection, and X and Y be random variables taking values in $\mathcal{S}$ and $\mathcal{R}$ such that $Y = f(X)$. Then the entropy of Y equals the entropy of X; i.e., $H(Y) = H(X)$.

Since the use of bijections is central to our arguments, we consider discrete-time, discrete-state analogs of the queueing systems $\cdot$/M/1 and $\cdot$/GI/1-LCFS. However, some queueing systems are either more easily studied in continuous time or have properties that require a continuous-time formulation. For example, one such property is that the superposition of two independent simple point processes is a simple point process. A "simple" point process is one that almost surely does not have more than one point occurring at the same time. Such a feature cannot be guaranteed when time is discrete. It is necessary to deal with superpositions when studying networks of queues, where the arrival process to a node can be the superposition of departures from other nodes. We do not attempt a study of continuous-time queueing systems in this paper.

The rest of this section introduces some notation that will be used subsequently. Sections II and III, respectively, establish entropy increasing properties for a FCFS queue with i.i.d. geometric service times, and for a pre-emptive resume LCFS queue dispensing i.i.d. service times. Section IV considers the timing capacity of an FCFS queue. The Appendix presents a derivation, in discrete time, of previously known results concerning the connection between the entropy rate of the time- and Palm-stationary versions of a point process. These results are needed for the proof of Theorem 2 in Section III.

Manuscript received October 12, 2000; revised September 9, 2002. The work of B. Prabhakar was supported in part by a Terman Fellowship and an Alfred P. Sloan Fellowship. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Washington, DC, June 2001.

B. Prabhakar is with the Departments of Electrical Engineering and Computer Science, Stanford University, Stanford, CA 94305 USA (e-mail: balaji@stanford.edu).

R. Gallager is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA (e-mail: gallager@lids.mit.edu).

Communicated by V. Anantharam, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2002.807287

A. Notation

The basic discrete-time queueing model is one in which arrivals take place just at the beginning of time slots and departures take place just before the end of time slots. Suppose that a_n is the arrival time of the nth packet to the queue. The numbering is such that $-\infty < \cdots < a_{-1} < a_0 < 0 \le a_1 < a_2 < \cdots < \infty$. We will implicitly assume throughout the paper that the numbers of packets are marked upon them. Let $A_n = a_{n+1} - a_n$ be the interarrival time between packets numbered n and $n+1$, and let $\boldsymbol{A}$ denote the process $\{A_n, n \in \mathbb{Z}\}$. Let s_n be the service time requirement of the nth packet. Denote by $\boldsymbol{S}$ the process $\{s_n, n \in \mathbb{Z}\}$. For stability, it is assumed that $E(s_1) < E(A_1)$.

II. Entropy and the ·/Geom/1 FCFS Queue

Consider a single server FCFS queue at which the services are independent and geometrically distributed with mean $1/\mu$. Specifically, let $P(s_0 = k) = \mu(1-\mu)^{k-1}$, for $k \ge 1$. Note that this disallows more than one packet from departing in a time slot. We shall also insist that $\mu < 1$, so that the service times are not exactly equal to 1 (if the service times equal 1 a.s., then no queues will form and the departure process is simply equal to the arrival process shifted by one unit of time).

This queue is the discrete-time analog of the exponential server queue, and we shall denote it symbolically as ·/Geom/1. The arrival process $\boldsymbol{A}$ is assumed to be stationary and ergodic, with $E(A_1) > E(s_1)$, and is independent of the service times $\{s_n, n \in \mathbb{Z}\}$. The waiting time of the nth packet may be obtained via Lindley's equation as follows:

$$w_n = \sup_{j \le n-1} \left\{ \sum_{i=j}^{n-1} (s_i - A_i), 0 \right\}. \tag{1}$$

As a result of the stability assumption ($E(s_1) < E(A_1)$), the random variables w_n are known to be finite a.s. (see Loynes [11]). The departure time of the nth packet may then be obtained from the equation

$$d_n = a_n \mathbb{1}_{\{w_n = 0\}} + d_{n-1} \mathbb{1}_{\{w_n > 0\}} + s_n. \tag{2}$$

Thus, (1) and (2) completely specify the departure times in terms of the arrival and service times. Let F_d be the function defined by (1) and (2) such that

$$d_n = F_d(s_n, s_{n-1}, \ldots; a_n, a_{n-1}, \ldots).$$

In the sequel, we will often deal with finite sequences of the form $\{s_k, a_k, l \le k \le n\}$. Given such a sequence *and* d_{l-1}, one can obtain the departure sequence $\{d_k, l \le k \le n\}$ recursively from the equation

$$d_k = \max\{a_k, d_{k-1}\} + s_k, \qquad l \le k \le n. \tag{3}$$

Let $\overline{F}_d^{\,n, l-1}$ denote the recursion

$$d_n = \overline{F}_d^{\,n, l-1}(s_n, \ldots, s_l; a_n, \ldots, a_l; d_{l-1}).$$

On the other hand, from the arrival and departure times one may deduce the service times using the equation

$$s_n = d_n - \max\{a_n, d_{n-1}\}. \tag{4}$$

Analogous to the recursions F_d and $\overline{F}_d^{\,n, l-1}$, (4) defines functions F_s and $\overline{F}_s^{\,n, l-1}$ such that

$$s_n = F_s(a_n, a_{n-1}, \ldots; d_n, d_{n-1}, \ldots)$$
$$s_n = \overline{F}_s^{\,n, l-1}(a_n, \ldots, a_l; d_n, \ldots, d_l; d_{l-1}).$$

It is well known [18] that if $\{A_n, n \in \mathbb{Z}\}$ is i.i.d., $P(A_1 = k) = \lambda(1-\lambda)^{k-1}$ for $k \ge 1$ and $\lambda < \mu$, then the inter-departure time sequence $\{D_n = d_{n+1} - d_n, n \in \mathbb{Z}\}$ is distributed identically as $\{A_n, n \in \mathbb{Z}\}$. For general stationary and ergodic arrival processes, given that $E(A_1) > E(s_1)$, the result of Loynes' [11] asserts that the departure process $\boldsymbol{D}$ is also stationary and ergodic with $E(D_1) = E(A_1)$ Indeed, this is easy to see from Lindley's recursions for the waiting and inter-departure times

$$w_{n+1} = [w_n + s_n - A_n]^+$$
$$D_n = [A_n - w_n - s_n]^+ + s_{n+1}.$$

The first equation and the joint stationarity and ergodicity of $\{(A_n, s_n), n \in \mathbb{Z}\}$ implies that $\{(A_n, s_n, w_n), n \in \mathbb{Z}\}$ is jointly stationary and ergodic. This and the second equation imply that $\{(A_n, s_n, w_n, D_n), n \in \mathbb{Z}\}$ is jointly stationary and ergodic. Unless explicitly stated otherwise, all queues and networks considered in this paper are assumed to be stable and in equilibrium.

Let $A^N = (A_1, \ldots, A_N)$, $A^{-N} = (A_{-N}, \ldots, A_{-1})$, and $A^{-N,N} = (A_{-N}, \ldots, A_N)$. If $H(A^N)$ is the entropy of A^N, then the *entropy rate* of $\boldsymbol{A}$ is defined as

$$H_{ER}(\boldsymbol{A}) = \lim_{N \to \infty} H(A^N)/N.$$

By the stationarity of the sequence $\{A_n, n \in \mathbb{Z}\}$, it follows that

$$H_{ER}(\boldsymbol{A}) = \lim_{N \to \infty} H(A^{-N})/N = \lim_{N \to \infty} H(A^{-N,N})/(2N+1).$$

A similar definition holds for $H_{ER}(\boldsymbol{D})$.

We are now ready to prove the following theorem which is the main result of this section.

Theorem 1: Let $\boldsymbol{A} = \{A_n, n \in \mathbb{Z}\}$ be an i.i.d. sequence of interarrival times with mean $1/\lambda$ according to which packets arrive at a ·/Geom/1 queue with service time equal to $1/\mu$, where $1 > \mu > \lambda$. Let $\boldsymbol{D} = \{D_n, n \in \mathbb{Z}\}$ be the corresponding interdeparture times. Then, $H_{ER}(\boldsymbol{A}) \le H_{ER}(\boldsymbol{D})$ with equality iff A_1 is geometric.

Proof: Consider the mutual information $I(A^N; D^N)$ between the vectors of the first N interarrival and interdeparture times. We may express it in the following two ways:

$$I(A^N; D^N) = H(A^N) - H(A^N | D^N) = H(D^N) - H(D^N | A^N).$$

This implies

$$H(D^N) - H(A^N) = H(D^N | A^N) - H(A^N | D^N). \tag{5}$$

Dividing by N and taking limits, we get

$$H_{ER}(\boldsymbol{D}) - H_{ER}(\boldsymbol{A}) = \lim_{N \to \infty} \frac{H(D^N | A^N) - H(A^N | D^N)}{N}.$$

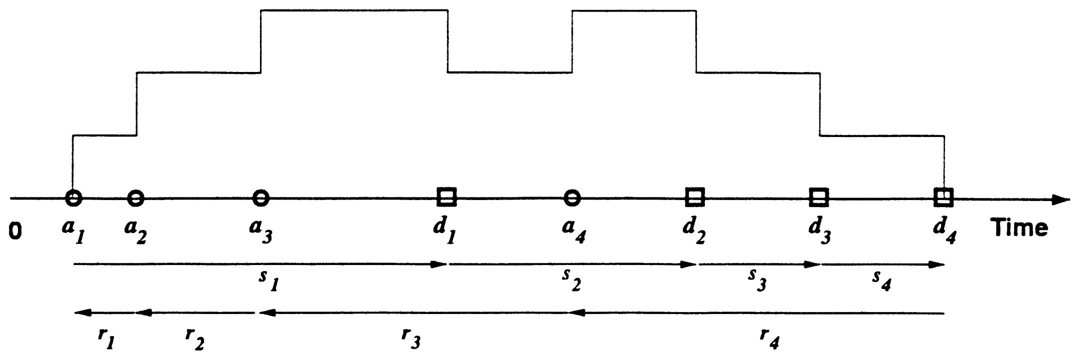

Fig. 1. A realization of the queue-size process of an FCFS queue. Arrival times are indicated by circles and departure times are indicated by squares.

Our method of proof will be to show that

$$\lim_{N\to\infty} \frac{H(D^N|A^N) - H(A^N|D^N)}{N} \geq 0$$

with equality iff A_1 is geometrically distributed.

Accordingly, first consider the term $H(D^N|A^N)$. From the relationships

$$a_k = a_1 + \sum_{i=1}^{k-1} A_i \quad \text{and} \quad d_k = d_1 + \sum_{i=1}^{k-1} D_i, \quad 2 \leq k \leq N+1$$

we see that there is a bijection between

$$(d_{N+1}, \ldots, d_1; a_{N+1}, \ldots, a_1) \quad \text{and} \quad (D^N; A^N; d_1, a_1).$$

We shall denote bijections symbolically as "$\leftrightarrow$." Thus,

$$(d_{N+1}, \ldots, d_1; a_{N+1}, \ldots, a_1) \leftrightarrow (D^N; A^N; d_1, a_1).$$

We wish to obtain a bijection involving the service times s_n. Observe from the relationships

$$d_k = \overline{F}_d^{k,0}(s_k, \ldots, s_1; a_k, \ldots, a_1; d_0), \quad \text{for } 1 \leq k \leq N+1$$

and

$$s_k = \overline{F}_s^{k,0}(a_k, \ldots, a_1; d_k, \ldots, d_1; d_0) \quad \text{for } 1 \leq k \leq N+1$$

that the vectors

$$(d_{N+1}, \ldots, d_1; a_{N+1}, \ldots, a_1; d_0)$$

and

$$(s_{N+1}, \ldots, s_1; a_{N+1}, \ldots, a_1; d_0)$$

uniquely specify each other through the functions $\overline{F}_d^{\cdot,0}$ and $\overline{F}_s^{\cdot,0}$. Thus, we have

$$(s_{N+1}, \ldots, s_1; A^N; a_1, d_0) \leftrightarrow (D^N; A^N; d_1, a_1, d_0). \quad (6)$$

We are now ready to express the term $H(D^N|A^N)$ in a form that is conducive to further analysis. Consider the following:

$$\begin{aligned} &H(D^N|A^N) \\ &= H(D^N; d_1, a_1, d_0|A^N) - H(d_1, a_1, d_0|A^N; D^N) \\ &\overset{(a)}{=} H(s_{N+1}, \ldots, s_1, a_1, d_0|A^N) - H(d_1, a_1, d_0|A^N; D^N) \\ &= H(s_{N+1}, \ldots, s_1|A^N) + H(a_1, d_0|A^N; s_{N+1}, \ldots, s_1) \\ &\quad - H(d_1, a_1, d_0|A^N; D^N) \\ &\overset{(b)}{=} (N+1)H(s_1) + H(a_1, d_0|A^N; s_{N+1}, \ldots, s_1) \\ &\quad - H(d_1, a_1, d_0|A^N; D^N). \end{aligned} \quad (7)$$

Equality (a) uses the bijection in (6) (recall that if X and Y take values in a finite or countable set and $X \leftrightarrow Y$, then $H(X) = H(Y)$). Equality (b) is a consequence of the service times being i.i.d., and independent of the arrival process. Dividing by N and taking limits we get

$$\lim_{N\to\infty} \frac{H(D^N|A^N)}{N} = H(s_1) \quad (8)$$

since the last two terms in (b) of (7) vanish in the limit of (8). To see this, note that $a_1 \leq A_0$, therefore $E(a_1) < \infty$. Hence, $H(a_1) \leq H(G) < \infty$, where G is a geometric random variable with mean $E(a_1)$. Now, $d_1 = a_1 + \text{sys}_1$, where $\text{sys}_1 = d_1 - a_1$ is the system time of packet 1. By Lemma 1, sys_1 is geometrically distributed. Therefore, $E(d_1)$ and, hence, $H(d_1)$ are finite. Similarly, one can show that $H(d_0) < \infty$.

Now consider the term $H(A^N|D^N)$. We shall deal with it by looking at the queue in reversed time.

We imagine that the queue-size process in reversed time corresponds to that of another queue whose arrival process is $\boldsymbol{D}$, in reverse; and whose departure process is $\boldsymbol{A}$, also suitably reversed. Thus, in reversed time, packet n "arrives" at time d_n and "departs" at time a_n. Observe that packet $n+1$ arrives *before* packet n in reversed time (see Fig. 1).

Corresponding to the operation of the queue in reversed time, we associate "reverse service times" $\boldsymbol{R} = \{r_n, n \in \mathbb{Z}\}$ with the packets as follows:

$$r_n = -(a_n - \min\{d_n, a_{n+1}\}) = \min\{d_n, a_{n+1}\} - a_n. \quad (9)$$

The similarity between (9) and (4) is clear once we interchange the role of arrivals and departures. The interpretation is that r_n is the service time of the nth packet "in reverse." That is, if packets arrive according to $\boldsymbol{D}$ reversed and depart according to $\boldsymbol{A}$ reversed, r_n is the amount of time packet n would have spent at the head of the queue. Fig. 1 illustrates this interpretation for a sample realization. We can rewrite (9) and express a_n in terms of d_n and r_n as

$$a_n = \min\{d_n, a_{n+1}\} - r_n. \quad (10)$$

It is clear that, analogous to the bijection in (6), (9) and (10) imply the following bijection:

$$\begin{aligned} &(r_1, \ldots, r_{N+1}; D^N; a_{N+2}, d_{N+1}) \\ &\qquad \leftrightarrow (A^N; D^N; a_{N+1}, a_{N+2}, d_{N+1}). \end{aligned} \quad (11)$$

The joint stationarity and ergodicity of $(\boldsymbol{A}, \boldsymbol{S})$ also implies the stationarity and ergodicity of the process $(\boldsymbol{A}, \boldsymbol{S}, \boldsymbol{D}, \boldsymbol{R})$. We have already argued the ergodicity of $\boldsymbol{D}$. The ergodicity of $\boldsymbol{R}$ follows from rewriting (9) as

$$r_n = \min\{d_n - a_n, a_{n+1} - a_n\} = \min\{w_n + s_n, A_n\} \quad (12)$$

since $d_n - a_n$ is the system time of packet n, equal to the sum of its waiting and service times.

Proceeding

$$\begin{aligned} &H(A^N|D^N) \\ &\overset{(a)}{=} H(A^{-N}|D^{-N}) \\ &= H(A^{-N}; a_0, a_1, d_0/D^{-N}) - H(a_0, a_1, d_0|D^{-N}; A^{-N}) \\ &\overset{(b)}{=} H(r_{-N}, \ldots, r_0, a_1, d_0|D^{-N}) \\ &\quad - H(a_0, a_1, d_0|D^{-N}; A^{-N}) \\ &\overset{(c)}{=} H(r_{-N}, \ldots, r_0|D^{-N}) + H(a_1, d_0|D^{-N}; r_{-N}, \ldots, r_0) \\ &\quad - H(a_0, a_1, d_0|D^{-N}; A^{-N}) \end{aligned} \quad (13)$$

where (a) is due to the joint stationarity of $(\boldsymbol{A}, \boldsymbol{D})$ and (b) is due to the bijection in (11). Dividing by N and taking limits, we obtain

$$\lim_{N\to\infty} \frac{H(A^N|D^N)}{N} = \lim_{N\to\infty} \frac{H(r_{-N}, \ldots, r_0|D^{-N})}{N}$$

since the last two terms of (c) vanish in the limit. By the stationarity of $(\boldsymbol{D}, \boldsymbol{R})$, we may rewrite the last expression as

$$\begin{aligned} \lim_{N\to\infty} \frac{H(A^N|D^N)}{N} &= \lim_{N\to\infty} \frac{H(r_1, \ldots, r_{N+1}|D^N)}{N} \\ &= \lim_{N\to\infty} \frac{H(r_1, \ldots, r_N|D^N)}{N} \\ &\triangleq H_{ER}(\boldsymbol{R}/\boldsymbol{D}). \end{aligned}$$

Thus, in order to prove Theorem 1, it suffices to show that

$$H(s_1) - H_{ER}(\boldsymbol{R}|\boldsymbol{D}) \geq 0$$

with equality iff the interarrival times are geometrically distributed. But $H_{ER}(\boldsymbol{R}/\boldsymbol{D}) \leq H_{ER}(\boldsymbol{R}) \leq H(r_1)$, as a consequence of unconditioning. Hence it is sufficient to show that

$$H(s_1) - H(r_1) \geq 0 \quad (14)$$

with equality iff A_1 is geometrically distributed.

Observe that the average service time of a packet is the same in forward and reversed time; that is, $E(s_1) = E(r_1)$. This follows immediately from two facts, which are easily verified. 1) Packet m begins a busy period, say B_f, and packet $m+n$ terminates it in forward time iff packet $m+n$ begins a busy period, say B_r, and packet m terminates it in reversed time. This implies that there are exactly the same number of packets in busy cycles B_f and B_r. 2) The lengths of B_f and B_r are identical. Since the length of a busy period is the sum of service times of the packets involved in that busy period, it follows from the two previously mentioned facts and the law of large numbers that $E(s_1) = E(r_1)$.

Since s_1 is geometrically distributed and the geometric distribution uniquely maximizes the entropy of all positive, integer-valued distributions of a given mean, it follows that $H(s_1) - H(r_1) \geq 0$. To complete the proof of Theorem 1, it suffices to show that r_1 is not geometrically distributed unless A_1 is geometrically distributed.

So, what is the distribution of r_1?

From (12) we know that

$$r_n = \min\{d_n - a_n, A_n\} = \min\{\text{sys}_n, A_n\}$$

where $\text{sys}_n = w_n + s_n$ equals the system time of packet n. A moment's reflection shows that r_n is precisely the amount of time that packet n spends at the *very back* of the queue evolving in forward time (see Fig. 1 for an illustration). By contrast, the service time of packet n is the amount of time it spends at the *very front* of the queue evolving in forward time. Proceeding

$$\begin{aligned} P(r_n > N) &= P(\min\{\text{sys}_n, A_n\} > N) \\ &\overset{(a)}{=} P(\text{sys}_n > N)P(A_n > N) \end{aligned} \quad (15)$$

where (a) follows from the fact that the system time of packet n depends only upon interarrival times $\{A_k, k < n\}$ and service times $\{s_k, k \leq n\}$, and that A_n is independent of all these variables (by the renewal assumption on the arrival process). We proceed with the following lemma which says that the system time of a typical packet in a GI/Geom/1 system is geometrically distributed.

Lemma 1: Let $\boldsymbol{A} = \{A_n, n \in \mathbb{Z}\}$ be an i.i.d., mean $1/\lambda$ interarrival sequence according to which packets arrive at a ·/Geom/1 queue with mean service time $1/\mu < 1/\lambda$. Then the system time of a packet is geometrically distributed.

Proof: Let X_n be the total number of packets in the queue immediately after the arrival of packet n, including packet n and the one in service. It is a well-known fact of continuous-time queueing theory (see, for example, [17, Sec. 8-6]) that the total number of packets in a stable GI/M/1 queueing system immediately after the arrival of packet n is a Markov chain with a geometric equilibrium distribution. Adapting the same argument to discrete renewal arrivals and i.i.d. geometric services is straightforward and implies that X_n is geometrically distributed.

The system time of packet n is, therefore, equal to $\sum_{i=1}^{X_n} Y_i$, where the Y_i are the service times of the packets found in the queue by packet n when it arrives (and this includes its own service time). But the Y_i are i.i.d. geometric with mean $1/\mu$ and independent of X_n. Being a geometric sum of geometric random variables (r.v.'s), the system time of packet n is geometrically distributed. □

Using the conclusion of Lemma 1 in (15) we get for every N that

$$P(r_n > N) = c^N P(A_n > N), \qquad \text{for some } c < 1. \quad (16)$$

It follows that r_n is geometric iff A_n is geometric, proving Theorem 1. □

Corollary 1: Let $\boldsymbol{A}$ be a mean $1/\lambda$ ergodic, stationary interarrival process to a ·/Geom/1 queue with mean service time $1/\mu$, where $\lambda < \mu$. Let $\boldsymbol{D}$ be the corresponding departure process. If the interarrival times have a tail that decays faster than a geometric of mean $1/\mu$, i.e., $P(A_n > N) < (1-\mu)^N$ for all N large enough, then $H_{ER}(\boldsymbol{D}) > H_{ER}(\boldsymbol{A})$.

Proof: From the proof of Theorem 1, we know r_n equals the minimum of the system time of packet n and A_n. We also know that it has mean $1/\mu$. Since $E(r_1) = E(s_1)$, it suffices to show that r_n is not geometrically distributed for this would imply that $H_{ER}(\boldsymbol{R}) < H_{ER}(\boldsymbol{S})$ and hence that $H_{ER}(\boldsymbol{A}/\boldsymbol{D}) < H_{ER}(\boldsymbol{D}/\boldsymbol{A})$. Now

$$P(r_n > N) = P(\min\{\text{sys}_n, A_n\} > N) \leq P(A_n > N).$$

Since $P(A_n > N) < (1-\mu)^N$ for N large enough, it follows r_n cannot be geometric with mean $1/\mu$. □

Suppose $\boldsymbol{G} = \{G_i, i \in \mathbb{Z}\}$ is an i.i.d. mean $1/\lambda$ geometrically distributed sequence. The process $\boldsymbol{G}$ will be called the Geometric arrival process. For each $N \geq 1$, let

$$P(A_1 = i_1, \ldots, A_N = i_N) = p_{i_1, \ldots, i_N}.$$

Then, the relative entropy between A^N and G^N is

$$D(A^N \| G^N) = \sum_{i_1, \ldots, i_N} p_{i_1, \ldots, i_N} \log \left(\frac{p_{i_1, \ldots, i_N}}{\prod_{k=1}^{N} \lambda(1-\lambda)^{i_k - 1}} \right).$$

Define the "relative entropy rate" between $\boldsymbol{A}$ and $\boldsymbol{G}$ as

$$D_{ER}(\boldsymbol{A} \| \boldsymbol{G}) = \lim_{N \to \infty} \frac{D(A^N \| G^N)}{N}.$$

Corollary 2: Consider queueing systems that satisfy the hypotheses of Theorem 1 and/or Corollary 1. Let $\boldsymbol{G} = \{G_i, i \in \mathbb{Z}\}$ be an i.i.d. mean $1/\lambda$ geometrically distributed sequence. Then, $D_{ER}(\boldsymbol{A} \| \boldsymbol{G}) \geq D_{ER}(\boldsymbol{D} \| \boldsymbol{G})$ with equality iff the interarrival times are i.i.d. geometric.

Proof: Now

$$\begin{aligned} D(A^N \| G^N) &= \sum_{i_1, \ldots, i_N} p_{i_1, \ldots, i_N} \log \left(\frac{p_{i_1, \ldots, i_N}}{\prod_{k=1}^{N} \lambda(1-\lambda)^{i_k - 1}} \right) \\ &= \sum_{i_1, \ldots, i_N} p_{i_1, \ldots, i_N} \log p_{i_1, \ldots, i_N} \\ &\quad - \sum_{i_1, \ldots, i_N} p_{i_1, \ldots, i_N} \log \left(\prod_{k=1}^{N} \lambda(1-\lambda)^{i_k - 1} \right) \\ &= -H(A^N) + N \log \frac{1-\lambda}{\lambda} - \frac{N \log(1-\lambda)}{\lambda}. \end{aligned}$$

Therefore,

$$D_{ER}(\boldsymbol{A} \| \boldsymbol{G}) = -H_{ER}(\boldsymbol{A}) + \log \frac{1-\lambda}{\lambda} - \frac{\log(1-\lambda)}{\lambda}$$

and similarly

$$D_{ER}(\boldsymbol{D} \| \boldsymbol{G}) = -H_{ER}(\boldsymbol{D}) + \log \frac{1-\lambda}{\lambda} - \frac{\log(1-\lambda)}{\lambda}.$$

From Theorem 1, $H_{ER}(\boldsymbol{A}) \leq H_{ER}(\boldsymbol{D})$ with equality iff $\boldsymbol{A}$ is geometric. Therefore, $D_{ER}(\boldsymbol{A} \| \boldsymbol{G}) \geq D_{ER}(\boldsymbol{D} \| \boldsymbol{G})$ with equality iff $\boldsymbol{A}$ is geometric. □

Corollary 3: Let $\boldsymbol{A}$ be a mean $1/\lambda$ ergodic, stationary arrival process to a ·/Geom/1 queue with mean service time $1/\mu$, and let $\boldsymbol{D}$ be the corresponding departure process. The following statements hold.

1) If $\boldsymbol{A}$ is the Geometric arrival process, then so is $\boldsymbol{D}$.

2) The only renewal arrival process that is a fixed point for the queue (i.e., $\boldsymbol{A} \stackrel{d}{=} \boldsymbol{D}$) is the Geometric arrival process.

Proof: Both statements follow from Theorem 1. □

Statement 1) is the discrete-time equivalent of Burke's theorem [5]. Statement 2) is an entropy proof of the uniqueness of the geometric arrival process as a fixed point for the ·/Geom/1 queue among all renewal arrival processes. This result is contained in Anantharam [1], who used a metric on arrival processes to show that the only stationary and ergodic fixed point for the ·/M/1 queue is the Poisson process.

One can view the ·/Geom/1 queue as a "Markov operator," producing a departure process distribution from an arrival process distribution. Speaking in this somewhat abstract fashion, passing an arbitrary arrival process through a series of i.i.d. ·/Geom/1 queues is like watching the evolution of a discrete-time Markov chain (see [14]). This makes possible connections with such notions of standard Markov chain theory as the existence of invariant distributions and relative entropy. If ν_n is the distribution of a Markov chain on a (countable) state space at the nth step and if ν is the corresponding invariant distribution, then it is well known that $D(\nu_n \| \nu)$ goes to zero as $n \to \infty$.

The existence of an invariant distribution for the ·/Geom/1 queue is, of course, well known and rederived in Corollary 3 of this paper using entropy arguments. It is none other than the geometric arrival process G. Just as in standard Markov chain theory one expects that $D_{ER}(\boldsymbol{D}_n \| \boldsymbol{G})$, the relative entropy of the departure process from the nth station of a series of i.i.d. ·/Geom/1 queues with respect to the invariant distribution $\boldsymbol{G}$, decreases to zero as n goes to infinity. Corollary 2 provides a partial answer in that it shows that $D_{ER}(\boldsymbol{D}_n \| \boldsymbol{G})$ is nonincreasing.

It has been established in [12] via coupling arguments that $\boldsymbol{D}_n$ converges in distribution the $\boldsymbol{G}$; thus, there is a non-entropy argument for the desired convergence.

III. The ·/GI/1-LCFS Queue

Consider a queue at which the service times $\{s_n\}$ are nonnegative integer-valued, i.i.d., arbitrarily distributed and have a mean equal to $1/\mu$. We will be interested in showing that the entropy rate of the departure process is no less than that of the arrival process. In general, point processes have two representations: the time and the Palm representations. In the time version, the point process is viewed as a time-stationary and ergodic process taking values in the space of random point measures (or Radon measures), while in the Palm version one considers the interoccurrence process as a stationary and ergodic process. Palm processes are obtained from the corresponding time processes by restricting to the event that a point occurred at the origin. In the previous section and in the rest of the paper, we consider entropy rates with respect to the Palm measure. However, in this section we shall find it useful to invoke the time-stationary representation in order to prove Theorem 2. The Appendix reviews the connection between time and Palm entropy rates; specifically, it shows that the time entropy rate equals λ times the Palm entropy rate. Thus, an increase in one implies an increase in the other.

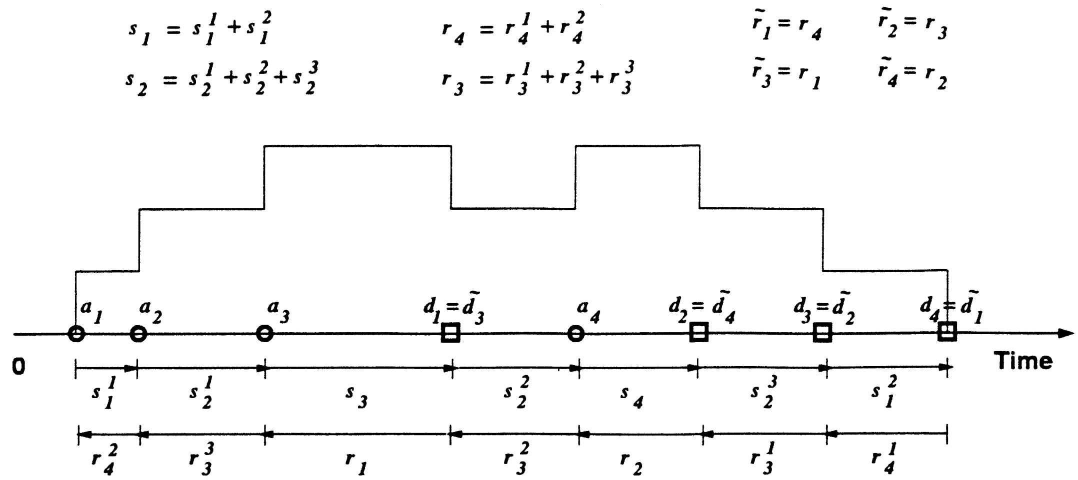

Fig. 2. A realization of a ./GI/1-LCFS queue.

We shall continue to assume that $P(s_1 = 0) = 0$. For technical reasons, we will also assume that the service times have a bounded support: that is, there is an $L < \infty$ such that $P(s_1 < L) = 1$. Note that this implies $H(s_1) \leq \log L$. To avoid trivialities, we will insist that $P(s_1 = 1) < 1$. Otherwise, every arriving packet receives only one unit of service and hence no queues can form. Suppose the service discipline is LCFS with preemptive resume. That is, every arriving packet p begins service immediately, interrupting any packet, say q, that may be in service, and q's service is resumed when p's service is completed. This service discipline is best visualized as a push-down stack, where arriving packets are placed at the top of the stack and the entire service effort is directed toward the top-most packet.

Fig. 2 shows a sample realization of a ·/GI/1-LCFS queue for positive time, assuming that there are zero packets queued at time 0.

We will continue to assume that packets are numbered according to their arrival times. Thus, packet n arrives at time a_n, where the a_ns are a strictly increasing sequence and $a_0 < 0 \leq a_1$. Let $\tilde{d}_n$ denote the departure time of packet n. As opposed to the FCFS service discipline, the departure order of packets may differ from their arrival order. Thus, although $\tilde{d}_n > a_n$ for each n, $\tilde{d}_n$ could be bigger than $\tilde{d}_{n+1}$. The reordering is illustrated in Fig. 2, where the packet arrival order is $1, 2, 3, 4$ while the departure order is $3, 4, 2, 1$.

For $n > 0$, let d_n be the *time of the* n*th departure* from the queue at or after time 0, and for $n \leq 0$, let d_n be the time of the $-(n-1)$th departure from the queue before time 0. Thus, $-\infty < \cdots < d_{-1} < d_0 < 0 \leq d_1 < \cdots < \infty$. For the rest of the paper, the departure process from the queue will be denoted by $\boldsymbol{D} = \{D_n = d_{n+1} - d_n,\, n \in \mathbb{Z}\}$. Note that with the definitions of arrival and departure times as above, if $a_n = d_k = M$, then the nth arrival occurs at time M^+ just after the kth departure has occurred at time M^-. This is a consequence of our assumption that arrivals take place at the beginning of time slots and departures take place at the end of time slots. Given the stationarity and ergodicity of the arrivals and service processes, it follows from Theorem 6 of the Appendix that the departure process is also stationary and ergodic (and thus its entropy rate is well defined).

It is a well-known fact of continuous-time queueing theory (see [10], [17], for example) that if the arrival process $\boldsymbol{A} = \{A_n,\, n \in \mathbb{Z}\}$ to a stable ·/GI/1-LCFS is i.i.d. exponential, the equilibrium departure process $\boldsymbol{D} = \{D_n,\, n \in \mathbb{Z}\}$ is also i.i.d. exponential. (Note that the departure process is defined in terms of the d_n's and not the $\tilde{d}_n$'s.) One expects an analogous result to be true in discrete time: When the services are i.i.d., nonnegative integer valued with a mean $1/\mu$ and when $\boldsymbol{A} = \{A_n,\, n \in \mathbb{Z}\}$ has i.i.d. geometric interarrival times with mean $1/\lambda$ (for $\lambda < \mu$), the departure process $\boldsymbol{D} = \{D_n,\, n \in \mathbb{Z}\}$ is also i.i.d. geometric. This follows as a simple corollary of Theorem 2 which shows that, subject to some restrictions on input and service distributions, the ·/GI/1-LCFS queue increases the entropy of a process passing through it.

As before, we will look at the queue in reversed time by changing the roles of the arrival and departure processes. Thus, in reversed time, arrivals occur according to $\boldsymbol{D}$ reversed and depart according to $\boldsymbol{A}$ reversed. The service discipline for the queue evolving in reversed time is also LCFS with preemption. We will again be interested in determining "reverse service times" for packets. Since packet departure orders may not be equal to packet arrival orders, we distinguish the reverse service time of the packet departing at time d_n and that of packet numbered n, which departs at time $\tilde{d}_n$. Accordingly, denote the former by r_n and the latter by $\tilde{r}_n$.

For convenience, we name the queue evolving in forward time Q_F and the queue evolving in reversed time Q_R. Because of the LCFS service policy, the forward service time s_n of packet n is precisely the total time it spends at the very back of Q_F. Since Q_R also employs the LCFS service policy, $\tilde{r}_n$ also equals the amount of time packet n spends at the very back of Q_R. A moment's reflection (aided by the illustration in Fig. 2) shows that packet n is at the back of Q_F precisely during the same instants of time that it is at the back of Q_R. Therefore, $\tilde{r}_n = s_n$.

We now relate the reverse service times $\tilde{r}_n$ and r_n. By definition, $\tilde{r}_n$ is the reverse service time of packet numbered n, whereas r_n is the reverse service time of the packet arriving to

Q_R at time d_n. Therefore, what is the number of the packet arriving to Q_R at time d_n? Equally, what is the number of the packet departing from Q_F at time d_n? Since only one packet departs from Q_F per time slot (recall $s_n \geq 1$ a.s.), there is a random one-to-one mapping, $T\colon \mathbb{Z} \to \mathbb{Z}$, taking packet arrival orders into packet departure orders. That is, if $T(k) = n$, then the kth arriving packet is the nth departing packet. Thus, the packet departing at time d_n is numbered $T^{-1}(n)$. This implies $r_n = \tilde{r}_{T^{-1}(n)} = s_{T^{-1}(n)}$.

As in the case of FCFS queues, showing that the departure process has a higher entropy rate than the arrival process reduces to showing that entropy rate of the forward service times is greater than the entropy rate of the reverse service times. However, the entropy rate of the reverse service times process, $\lim_n H(r_1, \ldots r_n)/n$, need not exist, since the process $\{r_n, n \in \mathbb{Z}\}$ is not stationary in general. This can be shown to be a consequence of the so-called "Inspection Paradox": The reverse service time of the first packet to depart after time 0 is likely to be longer than that of a typical departure.[1] We deal with this technicality as follows.

Let B_i be the busy cycles of Q_F, where B_0 is the busy cycle initiated just before (and possibly including) time 0. Suppose there are M_0 arrivals in B_0 *after* time 0 and, for $i \geq 1$, let $M_i = M_0$ plus the number of arrivals in busy cycles $B_1, \ldots, B_i$. Note that M_i increases to ∞.

Definition 1: Consider a stable discrete-time ·/GI/1-LCFS queue fed by stationary and ergodic arrival processes. It is said to satisfy the M-condition if there exists an $M < \infty$ such that

$$\lim_{i\to\infty} \frac{M_i}{i} = M \quad \text{a.s.}$$

Conditions: Reference [17, Theorem 9, p. 422] asserts that the M-condition is met if the arrivals are renewal. In fact, in this case, the family of random variables $\{M_i - M_{i-1}, i > 0\}$ will be i.i.d. with finite first moment.

Now, the numbers of the packets arriving from time 0 until the termination of $B_i, i \geq 1$, is the set $F_i = \{1, \ldots, M_i\}$. Let q be the number of packets in the queue at time 0^-. Thus, there are q partially processed packets in the queue at time 0 which arrived during negative time. From time 0 through the end of B_i, $i \geq 1$, there will be exactly $q + M_i$ departures. The numbers of these departing packets are in the set

$$G_i = \{T^{-1}(1), \ldots, T^{-1}(q + M_i)\}.$$

Consider the first M_i departures and let

$$H_i = \{T^{-1}(1), \ldots, T^{-1}(M_i)\}$$

be the set of associated packet numbers. Take any positive entry $k \in H_i$. This means packet k was among the first M_i departures after time 0. Since k is positive, packet k could have been among the first M_i departures only if it had been among the first M_i arrivals. Therefore, $k \in F_i$. This and the fact that F_i contains only positive entries implies that

$$F_i \cap H_i = \{k \geq 1, k \in H_i\}.$$

[1] We thank Venkat Anantharam for this observation, leading to the correction of a previous argument.

Now

$$\sum_{k=1}^{M_i} 1_{\{r_k=j\}} = \sum_{k=1}^{M_i} 1_{\{s_{T^{-1}(k)}=j\}} = \sum_{m\in H_i} 1_{\{s_m=j\}} = \sum_{m\in F_i\cap H_i} 1_{\{s_m=j\}} + \sum_{m\in H_i - F_i} 1_{\{s_m=j\}}$$

and

$$\begin{aligned}\sum_{k=1}^{M_i} 1_{\{s_k=j\}} - 1_{\{r_k=j\}} &= \sum_{k\in F_i\cap H_i} 1_{\{s_k=j\}} + \sum_{k\in F_i - H_i} 1_{\{s_k=j\}} \\ &\quad - \sum_{k\in F_i\cap H_i} 1_{\{s_k=j\}} - \sum_{k\in H_i - F_i} 1_{\{s_k=j\}} \\ &= \sum_{k\in F_i - H_i} 1_{\{s_k=j\}} - \sum_{k\in H_i - F_i} 1_{\{s_k=j\}}.\end{aligned}$$

Notice that the cardinality of the set $F_i - H_i$ is at most equal to q. This follows from the following facts: i) $T^{-1}(k) \leq M_i$ for $1 \leq k \leq M_i$ and ii) there are at most q negative entries in H_i. Also note that the cardinality of $H_i - F_i$ equals the cardinality of $F_i - H_i$. This gives

$$\frac{-q}{M_i} \leq \frac{\sum_{k=1}^{M_i} 1_{\{s_k=j\}} - 1_{\{r_k=j\}}}{M_i} \leq \frac{q}{M_i}.$$

As $i \to \infty$, $M_i \to \infty$, and

$$\lim_{i\to\infty} \frac{\sum_{k=1}^{M_i} 1_{\{s_k=j\}} - 1_{\{r_k=j\}}}{M_i} = 0.$$

For each n, there is an i such that $M_{i-1} \leq n \leq M_i$, and

$$\frac{-q - M_i + M_{i-1}}{M_i} \leq \frac{\sum_{k=1}^{n} 1_{\{s_k=j\}} - 1_{\{r_k=j\}}}{n} \leq \frac{q + M_i - M_{i-1}}{M_{i-1}}.$$

Since M_i/i converges a.s. to M, we get

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} 1_{\{s_k=j\}} - 1_{\{r_k=j\}}}{n} = 0 \quad \text{a.s.}$$

Further, since

$$\left| \frac{\sum_{k=1}^{n} 1_{\{s_k=j\}} - 1_{\{r_k=j\}}}{n} \right| \leq 2$$

by bounded convergence it follows that

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} P(s_k = j) - P(r_k = j)}{n} = 0. \tag{17}$$

Lemma 2: Consider a ·/GI/1-LCFS queue fed by an arbitrary stationary and ergodic arrival process. Suppose that the service

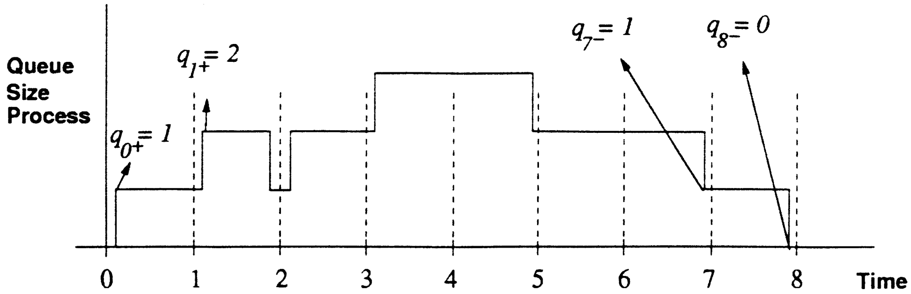

Fig. 3. The queue-size process.

times have a finite support; i.e., there is an $L < \infty$ such that $P(s_1 < L) = 1$. Then

$$\limsup_{n\to\infty} \frac{H(r_1, \ldots, r_n)}{n} \le H(s_1).$$

Proof: For each n define the Cesaro random variables C_n as

$$P(C_n = j) = \frac{\sum_{k=1}^{n} P(r_k = j)}{n}, \qquad \text{for } j = 0, 1, \ldots, L-1.$$

By concavity of the entropy function we get that $\frac{H(r_1, \ldots, r_n)}{n} \le H(C_n)$. We know from (17) that C_n converges in distribution to s_1. Since s_1, and hence all the C_n, have a bounded support, it follows that $H(C_n) \to H(s_1)$. This proves the lemma. □

Queue-Size Process: Before proceeding further we mention an important feature of discrete-time queues relevant for the main result of this section. As mentioned at the outset, we assume that arrivals occur just after the beginning of a time slot and departures occur just before the end of a time slot. Thus, the queue size q_k at any time k has two components: q_{k^-}, measured just after possible departures at the end of time $k-1$ and q_{k^+}, measured just after possible arrivals at the beginning of time k. This is illustrated in Fig. 3. Since $q_{k^+} - q_{(k+1)^-} \le 1$ it follows that $E(q_{k^-}) < \infty$ iff $E(q_{k^+}) < \infty$.

Definition 2: A stable discrete-time ·/GI/1-LCFS queue is said to satisfy the Q-condition if the number of packets in the queue in equilibrium has a finite first moment; i.e., $E(q_{0^-}) < \infty$.

Some Conditions: We list a few well-known necessary and sufficient conditions on the arrival and service processes for a ·/GI/1-LCFS queue to satisfy the Q-condition (for details, see [7]). Given that the services are i.i.d., it is necessary that $E(s_1^2) < \infty$. Each of the following conditions on the arrival process $\boldsymbol{A}$ is sufficient: $\boldsymbol{A}$ is i) renewal, or ii) strongly mixing. Thus, a wide variety of ·/GI/1-LCFS queues satisfy the Q-condition. We are interested in the Q-condition because of the following lemma.

Lemma 3: Let $\{q_k = (q_{k^-}, q_{k^+}), k \in \mathbb{Z}\}$ be the equilibrium queue-size process of a ·/GI/1-LCFS queue satisfying the Q-condition. Then, $H(q_0) < \infty$.

Proof: The random variables q_{k^-} and q_{k^+} are nonnegative, integer-valued, and have finite means. Their entropies are majorized by geometric random variables with means equal to $E(q_{k^-})$ and $E(q_{k^+})$, respectively. It follows that $H(q_{k^-}) + H(q_{k^+}) < \infty$. □

The M- and Q-conditions will be used in bijections related to LCFS queues. To relate the evolution of the queue in forward and reverse times we need to consider "attained" and "residual" service times (defined below). The M- and Q-conditions ensure that the entropy of these quantities is finite, allowing us to take limits. The details are made clear in Theorem 2. But, first, we define attained and residual service times.

Attained and Residual Services: Let u_{k^-} denote the ordered vector of packets in queue Q_F at time k^- together with the amount of service each has *already received*. Let v_{k^+} denote the ordered vector of packets in queue Q_F at k^+ along with the amount of service each has *yet to receive*.

Lemma 4: Consider an LCFS queue satisfying the M- and Q-conditions. If the support of the service times is bounded by L, then there is a $C < \infty$ not depending on k such that $H(v_{k^+}) + H(u_{k^-}) < C$.

Proof: Write $v_{k^+} = (X_1, \ldots, X_{q_{k^+}})$, where X_i is the service *yet to be received* by the ith packet from the front of the queue at time k and consider

$$\begin{aligned} H(v_{k^+}) &= H(X_1, \ldots, X_{q_{k^+}}) \\ &= H(q_{k^+}) + \sum_{i=0}^{\infty} P(q_{k^+} = i) H(X_1, \ldots, X_i | q_{k^+} = i) \\ &\le H(q_{k^+}) + \sum_{i=0}^{\infty} P(q_{k^+} = i) \sum_{j=1}^{i} H(X_j | q_{k^+} = i) \\ &\le H(q_{k^+}) + \log L \sum_{i=0}^{\infty} i P(q_{k^+} = i) \\ &\qquad\qquad (\text{since } 1 \le X_i < L) \\ &= H(q_{k^+}) + \log L E(q_{k^+}) = C^+ < \infty \end{aligned} \tag{18}$$

where C^+ is a constant not depending on k since $\{q_{k^+}, k \in \mathbb{Z}\}$ is a stationary process.

Similarly, to show that $H(u_{k^-}) \le C^- < \infty$, write $u_{k^-} = (Y_1, \ldots, Y_{q_{k^-}})$, where Y_i is the amount of service *already received* by the ith packet from the front and argue as above. Finally, letting $C = C^+ + C^-$ proves the lemma. □

Having established the preliminaries, we are ready to prove the following theorem.

Theorem 2: Let $\boldsymbol{A}$ be a stationary and ergodic arrival process with mean interarrival time $1/\lambda$ arriving at a ·/GI/1-LCFS queue with mean service time $1/\mu < 1/\lambda$. Suppose that the arrival and service process are such that the M- and Q-conditions are satisfied. Suppose also that the service times have a finite support. Let $\boldsymbol{D}$ be the corresponding departure process. Then, $H_{ER}(\boldsymbol{A}) \le H_{ER}(\boldsymbol{D})$.

Proof: For $K > 0$ consider the queue-size process restricted to $[0, K]$: $\{q_0, \ldots, q_K\}$. Let

$$N(K) = \max\{n: a_n \le K\}$$

be the number of arrivals in $[0, K]$. It is not hard to see that the following bijection holds:

$$(q_0, \ldots, q_K, v_{K^+}, a_1) \leftrightarrow \left(q_0, v_{0^+}, a_1, A^{N(K)-1}, S^{N(K)}\right). \tag{19}$$

Note that if $a_1 > K$ then $N(K) = 0$. In this case, it is to be understood that $A^{N(K)-1}$ and $S^{N(K)}$ are the empty vectors.

By Lemma 3, it follows that $H(q_0) < \infty$. This and the ergodicity of the process $\{q_k, k \in \mathbb{Z}\}$ imply that it has a finite entropy rate. We have also seen that both $H(v_{0^+})$ and $H(v_{K^+})$ are uniformly bounded.

Now taking entropies at (19), dividing both sides by K, and letting K go to infinity we get

$$\begin{aligned} H_{ER}(\boldsymbol{Q}) &\triangleq \lim_{K\to\infty} \frac{H(\{q_k; 0 \le k \le K\})}{K} \\ &= \lim_{K\to\infty} \frac{H\left(A^{N(K)-1}, S^{N(K)}\right)}{K}. \end{aligned} \tag{20}$$

Consider the term $H(A^{N(K)-1}, S^{N(K)})$. By assumption, the service process is independent of the arrival process and hence of $N(K)$, which is the number of arrivals in $[0, K]$. Therefore,

$$\begin{aligned} &H\left(A^{N(K)-1}, S^{N(K)}\right) \\ &\quad= H\left(A^{N(K)-1}\right) + H\left(S^{N(K)} \,\middle|\, A^{N(K)-1}\right) \\ &\quad= H\left(A^{N(K)-1}\right) + H\left(S^{N(K)} \,\middle|\, N(K)\right) \\ &\quad\overset{(a)}{=} H\left(A^{N(K)-1}\right) + \sum_{k=0}^{\infty} P(N(K)=k) H(S^k) \\ &\quad\overset{(b)}{=} H\left(A^{N(K)-1}\right) + H(s_1) \sum_{k=0}^{\infty} k P(N(K)=k) \\ &\quad= H\left(A^{N(K)-1}\right) + \lambda(K+1) H(s_1) \end{aligned}$$

where (a) uses the independence of the service process from $N(K)$ and (b) uses the fact that it is i.i.d. Therefore, (20) becomes

$$\begin{aligned} H_{ER}(\boldsymbol{Q}) &= \lambda H_{ER}(\boldsymbol{S}) + \lim_{K\to\infty} \frac{H\left(A^{N(K)-1}\right)}{K} \\ &= \lambda H_{ER}(\boldsymbol{S}) + \lim_{K\to\infty} \frac{H\left(A^{N(K)}\right)}{K}. \end{aligned} \tag{21}$$

Consider the term

$$\lim_{K\to\infty} \frac{H(A^{N(K)})}{K}.$$

By the well-known Shannon–McMillan–Breiman theorem

$$\lim_{K\to\infty} \frac{-\log p(A^K)}{K} = H_{ER}(\boldsymbol{A}) \quad \text{a.s. and in } L^1.$$

Since the random variables $N(K)$ *increase* to ∞, it follows that

$$\begin{aligned} \lim_{K\to\infty} \frac{-\log p\left(A^{N(K)}\right)}{K} &= \lim_{K\to\infty} \frac{-\log p\left(A^{N(K)}\right)}{N(K)} \frac{N(K)}{K} \\ &= \lambda H_{ER}(\boldsymbol{A}) \quad \text{a.s.} \end{aligned}$$

Given the above almost-sure convergence, if we could show that the random variables $\frac{-\log p(A_1, \ldots, A_{N(K)})}{K}$ are uniformly integrable, it follows that

$$\lim_{K\to\infty} \frac{H(A^{N(K)})}{K} = \lambda H_{ER}(\boldsymbol{A}).$$

But establishing uniform integrability is technically quite involved. Instead, we appeal to a result of Papangelou [13] (stated as Theorem 5 in the Appendix) and obtain via Corollary 5 that

$$\lim_{K\to\infty} \frac{H(A^{N(K)})}{K} = \lambda H_{ER}(\boldsymbol{A}).$$

Using this in (21) gives

$$H_{ER}(\boldsymbol{Q}) = \lambda(H_{ER}(\boldsymbol{S}) + H_{ER}(\boldsymbol{A})).$$

Let $M(K) = \max\{n: d_n \le K\}$ be the number of departures in $[0, K]$. Analogous to the bijection in (19) we obtain

$$\begin{aligned} (u_{0^-}, q_0, \ldots, q_K) &\leftrightarrow \left(u_{K^-}, q_K, d_1, \ldots, d_{M(K)}, R^{M(K)}\right) \\ &\leftrightarrow \left(u_{K^-}, q_K, d_1, D^{M(K)-1}, R^{M(K)}\right) \end{aligned}$$

by evolving the queue backward in time. Again, we adopt the convention that if $d_1 > K$ then $D^{M(K)-1}$ and $R^{M(K)}$ are the empty vectors. This bijection gives

$$\begin{aligned} H_{ER}(\boldsymbol{Q}) &= \lim_{K\to\infty} \frac{H(\{q_k; 0 \le k \le K\})}{K} \\ &= \lim_{K\to\infty} \frac{H\left(D^{M(K)}, R^{M(K)}\right)}{K} \\ &\le \limsup_{K\to\infty} \frac{H\left(D^{M(K)}\right)}{K} + \frac{H\left(R^{M(K)}\right)}{K} \\ &\overset{(a)}{=} \lambda H_{ER}(\boldsymbol{D}) + \limsup_{K\to\infty} \frac{H\left(R^{M(K)}\right)}{K} \end{aligned} \tag{22}$$

where (a) follows from the ergodicity of $\boldsymbol{D}$ and Corollary 5. Now consider

$$\begin{aligned} &\frac{H\left(R^{M(K)}\right)}{K} \\ &\quad= \frac{H(M(K))}{K} + \frac{H\left(R^{M(K)} | M(K)\right)}{K} \\ &\quad\le \frac{H(M(K))}{K} + \frac{\sum_{i=1}^{\infty} P(M(K)=i) H(r_1, \ldots, r_i)}{K} \\ &\quad\le \frac{\log(K+2)}{K} + \frac{\sum_{i=1}^{\infty} i P(M(K)=i) H(C_i)}{K} \end{aligned} \tag{23}$$

where $H(M(K)) \le \log(K+2)$ since $0 \le M(K) \le K+1$ (recall that there is at most one departure per time slot), and C_i is as defined in Lemma 2. Since the service times have a finite support, contained in $[0, L)$, it follows that $H(C_i) \le \log L$. We know from the proof of Lemma 2 that $H(C_i) \to H(s_1)$. Therefore, given a $\delta > 0$, we may choose an I such that

$$\sup_{\{i>I\}} H(C_i) \le H(s_1) + \delta.$$

Using all this we obtain

$$\begin{aligned}
&\limsup_{K\to\infty} \frac{H\left(R^{M(K)}\right)}{K} \\
&\le \limsup_{K\to\infty} \frac{\sum_{i=1}^{I} iP(M(K)=i)H(C_i)}{K} \\
&\quad + \limsup_{K\to\infty} \frac{(H(s_1)+\delta)\sum_{i=I+1}^{\infty} iP(M(K)=i)}{K} \\
&\le \limsup_{K\to\infty} \frac{(\log L)\, I^2 + (H(s_1)+\delta)E(M(K))}{K} \\
&= \lambda(H(s_1)+\delta)
\end{aligned} \tag{24}$$

since the average departure rate $E(M(K))/K$ equals the average arrival rate λ. Since δ is arbitrary, it follows that

$$\limsup_{K\to\infty} \frac{H(R^{M(K)})}{K} \le \lambda H(s_1).$$

Using this at (22), we get

$$\lambda(H_{ER}(\boldsymbol{D}) + H(s_1)) \ge H_{ER}(\boldsymbol{Q}) = \lambda(H_{ER}(\boldsymbol{A}) + H(s_1)).$$

Or, $H_{ER}(\boldsymbol{D}) - H_{ER}(\boldsymbol{A}) \ge 0$. This proves Theorem 2. □

Corollary 4: Let $\boldsymbol{A}$ be an arrival process with i.i.d. geometric interarrival times of mean $1/\lambda$ arriving at a ·/GI/1-LCFS queue with mean service time $1/\mu$, where $\mu > \lambda$. Also, suppose that the service times have a bounded support. Then the departure process $\boldsymbol{D}$ is distributed as $\boldsymbol{A}$.

Proof: This follows from Theorem 2 and two facts: 1) $\boldsymbol{D}$ is stationary and ergodic with mean interoccurrence time equal to $1/\lambda$, and 2) among stationary and ergodic processes with a given mean interoccurrence time the geometric process uniquely maximizes entropy rate. Thus, $H_{ER}(\boldsymbol{D}) = H_{ER}(\boldsymbol{A})$ and $\boldsymbol{D}$ is geometric. □

IV. The Timing Capacity of Single Server Queues

The paper of Anantharam and Verdú [2] considers a (continuous-time) queue as a channel through which a transmitter sends a message encoded in the arrival times of packets. The receiver decodes the message by observing the departure times of the packets. The randomness of the service times of the packets corrupts, or distorts, the original message embedded in the arrival times. Bedekar and Azizoğlu [4] extend the results of [2] to discrete-time queues and also study some variations involving multiple services per time slot.

In this section, we consider the timing capacity of ·/GI/1-FCFS queues and use our approach involving bijections to rederive some results from [2] and [4]. As this section is somewhat tangential to the rest of the paper, whose main focus is demonstrating the entropy increasing property of queueing systems, we shall consider "timing capacity" only in the sense of maximizing input–output mutual information.[2]

Consider a ·/GI/1-FCFS queue with arrival process $\boldsymbol{A}$ and departure process $\boldsymbol{D}$. For each $\lambda < \mu$, let

$$\tilde{C}(\lambda) = \sup_{\mathcal{L}(\boldsymbol{A})} \lim_{N\to\infty} \frac{I(A^N;\, D^N)}{N} \tag{25}$$

where the supremum is taken over the laws of rate λ input processes. In order for the limit to exist, we shall assume that the inputs $\boldsymbol{A}$ are stationary and ergodic. Note that as defined, $\tilde{C}(\lambda)$ can be thought of as the "timing capacity" in bits per arrival (assuming the base of the logarithm is 2). It is more natural to define the capacity as the amount of information that can be transmitted per unit time. The fact that there are λ arrivals per unit time (or time slot) on average motivates the following definition.

Definition 3: The timing capacity of a ·/GI/1-FCFS queue in bits per unit time is defined to be $C(\lambda) = \lambda\tilde{C}(\lambda)$ for $\tilde{C}(\lambda)$ as defined in (25).

Remark: When the service times are i.i.d., so long as the server conserves work and does not interrupt the service of a packet, it does not matter what the service discipline is: the timing capacity comes out to be same for all service disciplines and only depends on the service distribution (see [2] for an elaboration). We are making the FCFS discipline explicit in the above definition, since the LCFS results in this paper are derived for a preemptive resume discipline.

Denote by $\tilde{C}_{\text{geom}}(\lambda)$ and $C_{\text{geom}}(\lambda)$ the above quantities specialized to i.i.d., geometric service times. To employ the notation developed in the previous section, we shall suppose that the queue is initially in equilibrium.

Theorem 3: Consider a ·/Geom/1-FCFS queue with mean service rate μ. For each fixed $\lambda < \mu$

$$C_{\text{geom}}(\lambda) = \lambda(H(G_\lambda) - H(G_\mu))$$

where G_λ and G_μ are geometric random variables with means $1/\lambda$ and $1/\mu$, respectively. Thus, the capacity-achieving arrival process has i.i.d. geometrically distributed interarrival times of mean $1/\lambda$.

Proof: We shall first evaluate

$$\lim_{N\to\infty} \frac{I(A^N;\, D^N)}{N}$$

for an arbitrary arrival process $\boldsymbol{A}$. Since

$$I(A^N;\, D^N) = H(D^N) - H(D^N|A^N)$$

we obtain that

$$\lim_{N\to\infty} \frac{I(A^N;\, D^N)}{N} = H_{ER}(\boldsymbol{D}) - H_{ER}(\boldsymbol{D}|\boldsymbol{A}).$$

But, by (8), $H_{ER}(\boldsymbol{D}|\boldsymbol{A}) = H(s_1) = H(G_\mu)$ for G_μ as defined.

[2] As mentioned in [2], no previous theorem guarantees that maximizing input–output mutual information is equivalent to determining the timing capacity of queues. A more careful treatment would follow the formalism presented in [2].

Therefore, $C_{\text{geom}}(\lambda) = \lambda[\sup_{\mathcal{L}(\boldsymbol{A})}(H_{ER}(\boldsymbol{D}) - H(s_1))]$. Since arrivals and services are independent, the supremum over arrival processes is not affected by the $H(s_1)$ term, and thus one only seeks to maximize $H_{ER}(\boldsymbol{D})$ by a good choice of $\boldsymbol{A}$.

Corollary 3 implies that the supremum is achieved when $\boldsymbol{A}$ is a geometric process with mean $1/\lambda$, for then $\boldsymbol{D}$ is distributed as $\boldsymbol{A}$. Hence, $C_{\text{geom}}(\lambda) = \lambda(H(G_\lambda) - H(G_\mu))$ for G_λ and G_μ as stated in the theorem. □

It is shown in [4] that amongst discrete-time FCFS queues with a fixed mean service time, the ·/Geom/1 has the worst capacity. A similar result is obtained in [2] in the continuous-time setting. Moreover, [2, Theorem 5] presents an upper bound on the capacity of a continuous-time ·/GI/1-FCFS queue. For discrete-time queues, the capacity $C(\lambda)$ of a ·/GI/1-FCFS queue with service time s_1 satisfies the bound

$$C(\lambda) \le C_{\text{geom}}(\lambda) + \lambda D(s_1 \| G_\mu).$$

Formally, let $\boldsymbol{A}$, $\boldsymbol{S}$, and $\boldsymbol{D}$ be the arrival, service, and departure processes from a ·/GI/1-FCFS queue. Suppose that $\boldsymbol{A}$ is stationary and ergodic and that the queue is stable. Since the bijections used in Section II to obtain (8) did not rely on the services being geometric, (8) is valid for general i.i.d. services as well. Therefore,

$$\lim_{N\to\infty} \frac{I(A^N;\, D^N)}{N} = H_{ER}(\boldsymbol{D}) - H(s_1).$$

But the departure process has rate λ and its entropy rate is dominated by that of the geometric process of the same rate. We get

$$\begin{aligned}\lim_{N\to\infty} \frac{I(A^N;\, D^N)}{N} &\le H(G_\lambda) - H(s_1)\\ &= H(G_\lambda) - H(G_\mu) + H(G_\mu) - H(s_1).\end{aligned}$$

Since the above inequality is true for all arrival processes, taking the supremum over all rate λ arrival processes on the left-hand side, we get that

$$C(\lambda) \le C_{\text{geom}}(\lambda) + \lambda(H(G_\mu) - H(s_1)).$$

Let $p_i = P(s_1 = i)$. Now from

$$\begin{aligned}D(s_1\|G_\mu) &= \sum_{i\ge 1} p_i \log \frac{p_i}{\mu(1-\mu)^{i-1}}\\ &= \sum_{i\ge 1} p_i \left(\log p_i + \log\frac{1-\mu}{\mu} - i\log(1-\mu)\right)\\ &= -H(s_1) + \log\frac{1-\mu}{\mu} - \frac{\log(1-\mu)}{\mu}\\ &= -H(s_1) - \frac{\mu\log\mu + (1-\mu)\log(1-\mu)}{\mu}\\ &= -H(s_1) + H(G_\mu)\end{aligned}$$

we obtain that $C(\lambda) \le C_{\text{geom}}(\lambda) + \lambda D(s_1\|G_\mu)$ as announced.

V. Conclusion

We have considered two discrete queueing systems and have shown that they increase the entropy of a process passing through them. Some of the arguments used in the proofs are of general interest, possibly applicable elsewhere. While our method of proof in Section III requires the Q- and M-conditions, we believe Theorem 2 will hold under less restrictive assumptions on the arrival and service processes. Perhaps the biggest benefit of approaching queueing systems from an entropy standpoint is one of interpreting well-known queueing results (e.g., Burke's theorem) in a new way.

Appendix

The purpose of this appendix is to present some known results about point processes and their entropies, and the ergodicity of derived quantities like queue size and departure processes. These results are used in the proof of Theorem 2.

We shall be particularly interested in the connections between Palm and time-stationary probabilities, and hence in the connection between the entropy rate of the interoccurrence process (the entropy per arrival) and the entropy rate of the point process (the entropy per unit time). The results of interest to us—Theorem 5, Corollary 5, and Theorem 6—will be derived in the simpler discrete-time setting of this paper to keep it self-contained. The main references for this appendix are [3], [8], [13], and [15].

A. *Point Processes in Discrete Time*

A *counting measure* on $\mathbb{Z}$ is a measure m on $(\mathbb{Z}, \mathcal{B})$, where $\mathcal{B}$ denotes the discrete topology on $\mathbb{Z}$, such that

a) $m(C) \in \{0, 1, \ldots, \infty\}$ for all $C \in \mathcal{B}$,

b) $m([a, b]) < \infty$ for all bounded intervals $[a, b] \subset \mathbb{Z}$.

Let $\boldsymbol{M}$ be the set of all counting measures, m, on $\mathbb{Z}$. Endow $\boldsymbol{M}$ with the sigma field $\mathcal{M}$ generated by functions $m \to m(C)$, where C is a subset of $\mathbb{Z}$. The pair $(\boldsymbol{M}, \mathcal{M})$ is known as the *canonical space of point processes*; and a point process is thought of as a measurable mapping from some probability space $(\Omega, \mathcal{F}, P)$ into $(\boldsymbol{M}, \mathcal{M})$.

Definition 4: A point process $\boldsymbol{B}$ is an $\boldsymbol{M}$-valued random variable represented as

$$\boldsymbol{B} = \sum_{n=-\infty}^{\infty} \delta_{t_n} \tag{26}$$

where

$$\cdots < t_{-1} < t_0 < 0 \le t_1 < \cdots < t_n < t_{n+1} < \cdots \quad \text{a.s.}$$

and δ_x is the point mass at x.

The random variable t_n specifies the position of the nth point of $\boldsymbol{B}$. For any measurable function $f\colon \mathbb{Z} \to \mathbb{R}^+$

$$\int_{-\infty}^{\infty} f(x)\, d\boldsymbol{B} = \sum_{n=-\infty}^{\infty} f(t_n).$$

The *translation* or *shift operator* Θ_t operating on $\boldsymbol{M}$ is defined for each $t \in \mathbb{Z}$ by

$$\Theta_t m(C) = m(C+t), \qquad C \in \mathcal{B}.$$

The point process $\boldsymbol{B}$ is said to be *stationary* with respect to the family $\{\Theta_t,\, t \in \mathbb{Z}\}$, if for every $t \in \mathbb{Z}$ both $\boldsymbol{B}$ and $\Theta_t(\boldsymbol{B})$ have the same law. The set $G \in \mathcal{M}$ is said to be Θ_t*-invariant* if $\Theta_t G = G$ for all $t \in \mathbb{Z}$. $\boldsymbol{B}$ is said to be *ergodic* iff all Θ_t-invariant sets

$$I = \{\omega\colon \Theta_t \boldsymbol{B}(\omega) = \boldsymbol{B}(\omega) \text{ for all } t \in \mathbb{Z}\}$$

have probability 0 or 1.

We shall only be interested in stationary and ergodic point processes as we are primarily concerned with the equilibrium behavior of stable queues under stationary and ergodic inputs.

It is convenient to think of the point process $\boldsymbol{B}$ as the binary-valued process $\{b_n,\, n \in \mathbb{Z}\}$, where

$$b_n = \sum_{i=-\infty}^{\infty} \mathbb{1}_{\{t_i = n\}}.$$

Clearly, $\{b_n,\, n \in \mathbb{Z}\}$ is an ergodic stationary process since $\boldsymbol{B}$ is ergodic stationary.

Let $\{B_n = t_{n+1} - t_n,\, n \in \mathbb{Z}\}$ be the set of interoccurrence times of the point process $\boldsymbol{B}$. Let $N_B([0,\, t])$ be the number of points of $\boldsymbol{B}$ in the interval $[0,\, t]$. Then, $\lambda = E(N_B(\{0\}))$ is the average rate of $\boldsymbol{B}$. Note that $\lambda = P(b_0 = 1) = P(t_1 = 0)$, and that $1/\lambda = E(B_1)$.

B. Palm Probability and Entropy

The Palm probability P^0 of the point process $\boldsymbol{B}$ is defined on the probability space $(\Omega^0,\, \mathcal{F}^0)$, where $\Omega^0 = \Omega \cap \{t_1 = 0\}$ and $\mathcal{F}^0 = \mathcal{F} \cap \Omega^0$. Thus, the Palm probability is supported by those sequences which have a point at the origin. It is well known (see [15], or [3, Chapter 1]) that P^0 is the distribution of the process of interoccurrence times $\{B_n,\, n \in \mathbb{Z}\}$, while P is the distribution of the process $\{b_n,\, n \in \mathbb{Z}\}$. Further, the two probabilities P^0 and P are related by the expression

$$P^0(\cdot) = P(\cdot | b_0 = 1). \tag{27}$$

Theorem 4 [15, Theorem 3]: The process $\{b_n,\, n \in \mathbb{Z}\}$ is stationary and ergodic with respect to (w.r.t.) P iff the process $\{B_n,\, n \in \mathbb{Z}\}$ is stationary and ergodic w.r.t. P^0.

The entropy rate of the point process $\boldsymbol{B} = \{b_n,\, n \in \mathbb{Z}\}$ is

$$\begin{aligned} H_{ER}^T &= E_P(-\log P(b_0 | b_{-1},\, b_{-2},\, \ldots)) \\ &= \lim_{N\to\infty} \frac{E_P(-\log P(b_1,\, \ldots,\, b_N))}{N}. \end{aligned}$$

Thus, it is the amount of information per *unit time* and the T superscript emphasizes this point. The quantity

$$\begin{aligned} H_{ER}^O &= E_{P^0}(-\log P^0(B_0 | B_{-1},\, B_{-2},\, \ldots)) \\ &= \lim_{N\to\infty} \frac{E_{P^0}(-\log P^0(B_1,\, \ldots,\, B_N))}{N} \end{aligned}$$

is the amount of information per *occurrence*, emphasized by the O superscript.

The following theorem was originally proved by Papangelou [13] in the general (and quite technical) setting of continuous-time point processes. We state and prove it in the discrete-time setting of this paper.

Theorem 5 [13, Theorems 3 and 3a]: The entropy rates H_{ER}^T and H_{ER}^O defined above are related by the equation $H_{ER}^T = \lambda H_{ER}^O$.

Proof: Establishing the equality

$$\begin{aligned} &E_P(-\log P(b_0 | b_{-1},\, b_{-2},\, \ldots)) \\ &\qquad = \lambda E_{P^0}(-\log P^0(B_0 | B_{-1},\, B_{-2},\, \ldots)) \end{aligned} \tag{28}$$

essentially boils down to relating the probabilities of events under the measures P and P^0. We shall establish these relationships after making some definitions.

Let $\theta = -t_0$, $\overleftarrow{b} = (b_{-1}, b_{-2}, \ldots)$, and $\overleftarrow{B} = (B_{-1}, B_{-2}, \ldots)$. Note that the variables $\overleftarrow{b}$ and $(\theta,\, \overleftarrow{B})$ generate the same sigma algebra. Let $t \in \mathbb{Z}^+$, $x \in C_B = \mathbb{Z}^{+\mathbb{Z}^+}$, and $y \in C_b = \{0,\, 1\}^{\mathbb{Z}^+}$ be such that

$$\{\theta = t,\, \overleftarrow{B} \in dx\} = \{\overleftarrow{b} \in dy\}.$$

With these definitions, we claim the following equations hold:

$$\begin{aligned} P\left(b_0 = 1,\, \overleftarrow{b} \in dy\right) &= P\left(B_0 = t,\, \theta = t,\, \overleftarrow{B} \in dx\right) \\ &= \lambda P^0\left(B_0 = t,\, \overleftarrow{B} \in dx\right) \end{aligned} \tag{29}$$

$$\begin{aligned} P\left(b_0 = 0,\, \overleftarrow{b} \in dy\right) &= P\left(B_0 > t,\, \theta = t,\, \overleftarrow{B} \in dx\right) \\ &= \lambda P^0\left(B_0 > t,\, \overleftarrow{B} \in dx\right) \end{aligned} \tag{30}$$

$$P\left(b_0 = 1 \,\middle|\, \overleftarrow{b} = y\right) = \frac{P^0\left(B_0 = t \,\middle|\, \overleftarrow{B} = x\right)}{P^0\left(B_0 \geq t \,\middle|\, \overleftarrow{B} = x\right)} \tag{31}$$

$$P\left(b_0 = 0 \,\middle|\, \overleftarrow{b} = y\right) = \frac{P^0\left(B_0 > t \,\middle|\, \overleftarrow{B} = x\right)}{P^0\left(B_0 \geq t \,\middle|\, \overleftarrow{B} = x\right)}. \tag{32}$$

We establish (29) and (31), the proofs of (30) and (32) are identical.

The first equation in (29) is immediate from the definition of t, x, and y. The second equation follows from

$$\begin{aligned} &P\left(B_0 = t,\, \theta = t,\, \overleftarrow{B} \in dx\right) \\ &\quad = P\left(B_0 = t,\, b_0 = 1,\, \overleftarrow{B} \in dx\right) \\ &\quad = P\left(B_0 = t,\, \overleftarrow{B} \in dx \,\middle|\, b_0 = 1\right) P(b_0 = 1) \\ &\quad = P^0\left(B_0 = t,\, \overleftarrow{B} \in dx\right) \lambda. \end{aligned}$$

Next consider (31)

$$\begin{aligned} P\left(b_0 = 1 \,\middle|\, \overleftarrow{b} = y\right) &= \frac{P\left(b_0 = 1,\, \overleftarrow{b} \in dy\right)}{P\left(\overleftarrow{b} \in dy\right)} \\ &= \frac{P\left(B_0 = t,\, \theta = t,\, \overleftarrow{B} \in dx\right)}{P\left(\theta = t,\, \overleftarrow{B} \in dx\right)}. \end{aligned} \tag{33}$$

From (29), the numerator equals $\lambda P^0(B_0 = t,\, \overleftarrow{B} \in dx)$. As for the denominator

$$\begin{aligned}
&P\left(\theta = t,\, \overleftarrow{B} \in dx\right)\\
&= P\left(b_{-1} = 0,\, b_{-2} = 0,\, \ldots,\, b_{-t} = 1,\, \overleftarrow{b}_{-t-1} \in dy\right)\\
&\overset{(a)}{=} P\left(b_0 = 1,\, b_1 = 0,\, \ldots,\, b_{t-1} = 0,\, \overleftarrow{b}_{-1} \in dy\right)\\
&= P(b_0 = 1)\\
&\quad \cdot P\left(b_1 = 0,\, \ldots,\, b_{t-1} = 0,\, \overleftarrow{b}_{-1} \in dy \,\middle|\, b_0 = 1\right)\\
&= \lambda P^0\left(B_1 \geq t,\, \overleftarrow{B}_0 \in dx\right)\\
&\overset{(b)}{=} \lambda P^0\left(B_0 \geq t,\, \overleftarrow{B}_{-1} \in dx\right)\\
&= \lambda P^0\left(B_0 \geq t,\, \overleftarrow{B} \in dx\right),\\
&\qquad \left(\text{since, by definition, } \overleftarrow{B} = \overleftarrow{B}_{-1}\right)
\end{aligned}$$

where (a) is due to the stationarity of $\{b_n,\, n \in \mathbb{Z}\}$ w.r.t. P and (b) is due to the stationarity of $\{B_n,\, n \in \mathbb{Z}\}$ w.r.t. P^0 (see Theorem 4). Using all this at (33) we get

$$\begin{aligned}
P\left(b_0 = 1 \,\middle|\, \overleftarrow{b} = y\right) &= \frac{P^0\left(B_0 = t,\, \overleftarrow{B} \in dx\right)}{P^0\left(B_0 \geq t,\, \overleftarrow{B} \in dx\right)}\\
&= \frac{P^0\left(B_0 = t \,\middle|\, \overleftarrow{B} = x\right)}{P^0\left(B_0 \geq t \,\middle|\, \overleftarrow{B} = x\right)}
\end{aligned}$$

thus establishing (31).

To economize on space, we compactify notation and set

$$\begin{aligned}
P^0\left(B_0 = t,\, \overleftarrow{B} \in dx\right) &= P^0(t,\, x)\\
P^0\left(B_0 = t \,\middle|\, \overleftarrow{B} = x\right) &= P^0(t|x)\\
P^0\left(B_0 > t,\, \overleftarrow{B} \in dx\right) &= P^0({}^{>}t,\, x)\\
P^0\left(B_0 > t \,\middle|\, \overleftarrow{B} = x\right) &= P^0({}^{>}t|x)\\
P^0\left(B_0 \geq t,\, \overleftarrow{B} \in dx\right) &= P^0({}^{\geq}t,\, x)\\
P^0\left(B_0 \geq t \,\middle|\, \overleftarrow{B} = x\right) &= P^0({}^{\geq}t|x).
\end{aligned}$$

Continuing with the proof of Theorem 5

$$\begin{aligned}
&H_{ER}^T\\
&= E_P(-\log P(b_0|b_{-1},\, b_{-2},\, \ldots))\\
&= -\int_{C_b} P\left(b_0 = 1,\, \overleftarrow{b} \in dy\right) \log P\left(b_0 = 1 \,\middle|\, \overleftarrow{b} \in dy\right)\\
&\quad - \int_{C_b} P\left(b_0 = 0,\, \overleftarrow{b} \in dy\right) \log P\left(b_0 = 0 \,\middle|\, \overleftarrow{b} \in dy\right).
\end{aligned}$$

Substituting the Palm probabilities for the various events in the last expression yields

$$\begin{aligned}
H_{ER}^T &= -\lambda \int_{C_B} \sum_{t=1}^{\infty} P^0(t,\, x) \log\left(\frac{P^0(t|x)}{P^0({}^{\geq}t|x)}\right)\\
&\quad - \lambda \int_{C_B} \sum_{t=1}^{\infty} P^0({}^{>}t,\, x) \log\left(\frac{P^0({}^{>}t|x)}{P^0({}^{\geq}t|x)}\right)\\
&= -\lambda \int_{C_B} \sum_{t=1}^{\infty} P^0(t,\, x) \log P^0(t|x) \qquad (34)\\
&\quad - \lambda \int_{C_B} \sum_{t=1}^{\infty} P^0({}^{>}t,\, x) \log P^0({}^{>}t|x) \qquad (35)\\
&\quad + \lambda \int_{C_B} \sum_{t=1}^{\infty} P^0({}^{\geq}t,\, x) \log P^0({}^{\geq}t|x). \qquad (36)
\end{aligned}$$

The expression at (34) equals λH_{ER}^O, while (35) and (36) cancel telescopically to yield

$$\lambda \int_{C_B} P^0\left(B_0 \geq 1,\, \overleftarrow{B} \in dx\right) \log P^0\left(B_0 \geq 1 \,\middle|\, \overleftarrow{B} = x\right)$$

which equals 0 since $P^0(B_0 \geq 1|\overleftarrow{B} = x) = 1$ for every $x \in C_B$. Therefore, $H_{ER}^T = \lambda H_{ER}^O$ and Theorem 5 is proved. □

Corollary 5: Consider the processes $\boldsymbol{B}$ and $\{B_n,\, n \in \mathbb{Z}\}$ defined above. Let $N(K)$ be the number of points of $\boldsymbol{B}$ in $[0,\, K]$. Then

$$\lim_{K \to \infty} \frac{H\left(B^{N(K)}\right)}{K} = \lambda H_{ER}^O(\boldsymbol{B}).$$

Proof: Observe that

$$(t_1,\, B_1,\, \ldots,\, B_{N(K)}) \leftrightarrow (b_0,\, \ldots,\, b_K).$$

Therefore,

$$\begin{aligned}
H\left(t_1,\, B_1,\, \ldots,\, B_{N(K)}\right) &= H\left(B^{N(K)}\right) + H\left(t_1 \,\middle|\, B^{N(K)}\right)\\
&= H(b_0,\, \ldots,\, b_K),
\end{aligned}$$

from which and Theorem 5 the corollary follows. □

C. Marked Point Processes

Stationary marked point processes and their associated Palm theory can be found in [3, Sec. 1.3]. We will only recall the bare essentials here.

Let

$$\boldsymbol{B} = \sum_{n=-\infty}^{\infty} \delta_{t_n} \qquad (37)$$

be a stationary and ergodic point process, and let $\{s_n,\, n \in \mathbb{Z}\}$ be an i.i.d. sequence of positive integer-valued random variables, independent of $\{t_n,\, n \in \mathbb{Z}\}$. We imagine t_n to be the arrival time of nth packet and s_n to be its service time. The process $\{b_n,\, n \in \mathbb{Z}\}$ defined by

$$b_n = \sum_{i=-\infty}^{\infty} s_i \mathbb{1}_{\{t_i = n\}}$$

is a marked point process which takes the value zero where there are no points of $\boldsymbol{B}$ and at the points of $\boldsymbol{B}$ it takes the value of the marks s_n. It is important to note that since the s_n can only take on positive integer values, this definition is unambiguous.

A marked point process is said to be stationary and ergodic if $\{b_n, n \in \mathbb{Z}\}$ defined above is stationary and ergodic w.r.t. P under time translations: that is, for each $k \in \mathbb{Z}$

$$\{b_n, n \in \mathbb{Z}\} \stackrel{d}{=} \{b_{n+k}, n \in \mathbb{Z}\}.$$

The stationarity and ergodicity of $\{b_n, n \in \mathbb{Z}\}$ follows from that of the arrival point process $\boldsymbol{B}$ and the fact that $\{s_n, n \in \mathbb{Z}\}$ is i.i.d. and independent of the arrival times $\{t_n, n \in \mathbb{Z}\}$ (see [3, Example 1.3.4]).

Theorem 6: Consider a $\cdot$/GI/1-LCFS queue with a stationary and ergodic arrivals process $\boldsymbol{B}$ defined as in (37) and service times $\{s_n, n \in \mathbb{Z}\}$. Suppose $E(s_n) < E(t_n - t_{n-1})$. Let $q_n = (q_{n^-}, q_{n^+})$ be the resulting equilibrium queue-size process, and let $d_n = \mathbb{1}_{\{q_{(n-1)^+} - q_{n^-} = 1\}}$ be the corresponding departure process. Then $\{q_n, n \in \mathbb{Z}\}$ and $\{d_n, n \in \mathbb{Z}\}$ are both stationary and ergodic with respect to time translations.

Proof: Let $\mathcal{Q} \subset \mathbb{Z}^{\mathbb{Z}} \times \mathbb{Z}^{\mathbb{Z}}$ and $\mathcal{B} \subset \mathbb{Z}^{\mathbb{Z}}$ be such that

$$\{q_n, n \in \mathbb{Z}\} \in \mathcal{Q} \Leftrightarrow \{b_n, n \in \mathbb{Z}\} \in \mathcal{B}.$$

The existence of sets $\mathcal{Q}$ and $\mathcal{B}$ as above is ensured by Loynes' pathwise construction of the queue-size process (given the LCFS service discipline) from the arrival and service processes. The equation

$$\begin{aligned} P(\{q_n, n \in \mathbb{Z}\} \in \mathcal{Q}) &= P(\{b_n, n \in \mathbb{Z}\} \in \mathcal{B}) \\ &= P(\{b_{n+k}, n \in \mathbb{Z}\} \in \mathcal{B}) \\ &= P(\{q_{n+k}, n \in \mathbb{Z}\} \in \mathcal{Q}) \end{aligned}$$

verifies the stationarity of the queue-size process. If $\mathcal{Q}$ is a shift-invariant event for the queue-size process then $\mathcal{B}$ is a shift-invariant event for $\{b_n, n \in \mathbb{Z}\}$. The ergodicity of the latter implies that $P(\mathcal{B}) = 0$ or 1. Therefore, $P(\mathcal{Q}) = 0$ or 1. This proves the ergodicity of the queue-size process.

The stationarity and ergodicity of the departure process follows immediately from that of the queue-size process. □

ACKNOWLEDGMENT

B. Prabhakar thanks C. Nair for a thorough reading of the paper and for useful feedback.

REFERENCES

[1] V. Anantharam, "Uniqueness of stationary ergodic fixed point for a ·/M/k node," *Ann. Appl. Probab.*, vol. 3, no. 1, pp. 154–173, 1993.
[2] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inform. Theory*, vol. 42, pp. 4–18, Jan. 1996.
[3] F. Baccelli and P. Brémaud, *Elements of Queueing Theory*. Berlin, Germany: Springer-Verlag, 1994.
[4] A. S. Bedekar and M. Azizoğlu, "The information-theoretic capacity of discrete-time queues," *IEEE Trans. Inform. Theory*, vol. 44, pp. 446–461, 1998.
[5] P. Burke, "The output of a queueing system," *Oper. Res.*, vol. 4, pp. 699–704, 1956.
[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York: Wiley, 1991.
[7] D. J. Daley and T. Rolski, "Finiteness of waiting-time moments in general stationary single-server queues," *Ann. Appl. Probab.*, vol. 2, no. 1, pp. 987–1008, 1992.
[8] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*. New York: Springer, 1988, Springer Series in Statistics.
[9] P. Glynn, unpublished notes.
[10] F. P. Kelly, *Reversibility and Stochastic Networks*. London, U.K.: Wiley, 1979.
[11] R. M. Loynes, "The stability of a queue with nonindependent inter-arrival and service times," *Proc. Camb. Philos. Soc.*, vol. 58, pp. 497–520, 1962.
[12] T. S. Mountford and B. Prabhakar, "On the weak convergence of departures from an infinite sequence of ·/M/1 queues," *Ann. Appl. Probab.*, vol. 5, no. 1, pp. 121–127, 1995.
[13] F. Papangelou, "On the entropy rate of stationary point processes and its discrete approximation," *Z. Wahrs.*, vol. 44, pp. 191–211, 1978.
[14] B. Prabhakar and N. Bambos, "Entropy methods for high speed communications," in *Proc. 1995 Conf. Information Sciences and Systems*. Baltimore, MD, 1995.
[15] C. Ryll-Nardzewski, "Remarks on processes of calls," in *Proc. 4th Berkeley Symp.*, vol. 2, Univ. Calif., Berkeley, 1961, pp. 455–465.
[16] J. A. Thomas, "On the Shannon capacity of discrete-time queues," in *Proc. 1997 Int. Symp. Information Theory*, 1997.
[17] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[18] M. E. Woodward, *Communication and Computer Networks: Modeling With Discrete-Time Queues*. Los Alamitos, CA: IEEE Computer Soc., 1994.