

ASYMPTOTICS FOR OPEN-LOOP WINDOW FLOW CONTROL

ARTHUR W. BERGER

*AT&T Bell Laboratories
Room HO-3H-601
Holmdel, NJ 07730-3030.*

WARD WHITT

*AT&T Bell Laboratories
Room MH-2C-178
Murray Hill, NJ 07974-0636.*

(Received May 1993; revised October 1993)

ABSTRACT

An open-loop window flow-control scheme regulates the flow into a system by allowing at most a specified window size W of flow in any interval of length L . The *sliding window* considers all subintervals of length L , while the *jumping window* considers consecutive disjoint intervals of length L . To better understand how these window control schemes perform for stationary sources, we describe for a large class of stochastic input processes the asymptotic behavior of the maximum flow in such window intervals over a time interval $[0, T]$ as T and L get large, with T substantially bigger than L . We use strong approximations to show that when $T \gg L \gg \log T$ an invariance principle holds, so that the asymptotic behavior depends on the stochastic input process only via its rate and asymptotic variability parameters. In considerable generality, the sliding and jumping windows are asymptotically equivalent. We also develop an approximate relation between the two maximum window sizes. We apply the asymptotic results to develop approximations for the means and standard deviations of the two maximum window contents. We apply computer simulation to evaluate and refine these approximations.

Key words: Open-loop Control, Traffic Descriptor, Sliding Window, Jumping Window, Scan Statistic, Strong Approximations, Extreme Values, Asymptotics.

AMS (MOS) subject classifications: 60F05, 60F15.

1. Introduction

The Hungarian influence on probability theory, and mathematics more generally, seems to be greater than can be accounted for solely by chance. In this paper we pay tribute to an outstanding example (who we also claim as an American), Lajos Takács, on the occasion of his 70th birthday, by applying some Hungarian probability theory to a problem of interest in the

design of emerging high-speed communication networks. In particular, we apply strong approximations as in Komlós, Major and Tusnády [9], [12], Csörgő and Révész [6] and Csörgő, Horváth and Steinebach [5] to study open-loop window flow-control schemes. We also refer to related work by Erdős and Rényi [7].

The most familiar window flow-control scheme is a closed-loop control mechanism. Sources regulate their flow into the system (e.g., packets into a network) by keeping track of the flow that has been transmitted but not yet received, typically using acknowledgements sent back from the receiver to the source. The source stops sending when the unacknowledged flow reaches a specified window size; e.g., see pp. 97 and 429 of Bertsekas and Gallager [2].

In contrast, open-loop window flow-control schemes use only local information. For high-speed communication networks, there is growing interest in open-loop control schemes, because end-to-end delays over long distances can become very large relative to the time significant change in the network state occurs; see Budka [3], Reibman and Berger [14] and Sections 2.2 and 3.3 of Roberts [15]. The idea in the open-loop window flow control is to simply limit the maximum input in any interval (window) of specified length. Our purpose here is to gain a better understanding of the way these open-loop window control schemes perform for stationary sources. To do this, we study open-loop sliding and jumping window control schemes via asymptotics.

Let $I(t)$ represent the *input* to a system in the interval $[0, t]$. We regard $I(t)$ as random, so that $I \equiv \{I(t) : t \geq 0\}$ is a real-valued stochastic process. We think of I as being nondecreasing and integer valued, but neither of these properties are required. Let T be the *time period* and let L be the *window length*. For simplicity, assume that T is a multiple of L . Let $J \equiv J(L, T, I)$ be the random variable representing the *maximum jumping window content* as a function of L , T , and I , i.e.,

$$J \equiv J(L, T, I) = \max\{I(kL) - I((k-1)L) : 1 \leq k \leq T/L\},$$

and let $S \equiv S(L, T, I)$ be the random variable representing the *maximum sliding window content* as a function of L , T , and I , i.e.,

$$S \equiv S(L, T, I) = \sup\{I(t+L) - I(t) : 0 \leq t \leq T-L\}.$$

In the statistics literature, the sliding window is also called the *scan statistic*; see Naus [12], [13]. More related literature is cited there.

A trivial consequence of these definitions is the ordering

$$J(L, T, I) \leq S(L, T, I) \text{ w.p.1.}$$

Moreover, since any interval of length L yielding a maximum for S is contained in two consecutive intervals in J , and since the maximum jumping window is at least as big as each of these,

$$S(L, T, I) \leq 2J(L, T, I) \text{ w.p.1.}$$

Hence, for any process I and any time interval $[0, T]$, the ratio J/S is a random variable with $0.5 \leq J/S \leq 1$. One purpose here is to develop approximations for this ratio.

In general, the distributions of J and S depend on the distribution of the stochastic process I in a rather complicated way. However, as T and L become large, some statistical regularity emerges. Since it is natural to consider large values of T and L in communication network

applications, in the remaining sections we discuss asymptotics. We start in §2 by reviewing the central limit theorem and strong approximations, which justify an approximation by Brownian motion. In §3 we consider extreme-value asymptotics for Brownian motion as $T \rightarrow \infty$. In §4 we provide conditions for the Brownian approximation to be asymptotically correct and establish limits for S and J as both T and L get large. In particular, there we develop support for the following relatively simple approximations for the random variables (and their means) provided that $\log T \ll L \ll T$, where \ll means much smaller than:

$$S \approx ES \approx \lambda L + \Phi_S(T, L) \sqrt{\lambda c^2 L}, \tag{1}$$

where

$$\Phi_S(T, L) = \sqrt{2 \log(T/L) + 2 \log \log T} - \frac{1.28}{\sqrt{2 \log(T/L)}};$$

$$J \approx EJ \approx \lambda L + \Phi_J(T, L) \sqrt{\lambda c^2 L}, \tag{2}$$

where

$$\Phi_J(T, L) = \sqrt{2 \log(T/L)} - \frac{\log \log(T/L) + 1.38}{2 \sqrt{2 \log(T/L)}};$$

and

$$\frac{J - \lambda L}{S - \lambda L} \approx \frac{EJ - \lambda L}{ES - \lambda L} \approx E \left[\frac{J - \lambda L}{S - \lambda L} \right] \approx \frac{\Phi_J(T, L)}{\Phi_S(T, L)}, \tag{3}$$

for Φ_S and Φ_J defined above. In (1)–(3), λ is the rate and c^2 is the asymptotic variability parameter of I (defined in (6) below). In (1)–(3) and throughout the paper, \log means the natural logarithm, i.e., to the base e . Without loss of generality (by the choice of the measuring units), we can let $\lambda = 1$. Hence, there are really only three parameters in (1)–(3): L , T , and c^2 .

The approximations for S and J in (1) and (2) have a simple interpretation: The first term λL is the approximate mean for a single interval of length L ; the term $\sqrt{\lambda c^2 L}$ is the approximate standard deviation for a single interval of length L ; and the remaining factors $\Phi_S(T, L)$ for S in (1) and $\Phi_J(T, L)$ for J in (2) represent the increase due to the maximization. It is significant that the factors Φ_S and Φ_J depend only on T and L , and not upon λ and c^2 .

Note that, by (3), the ratio $(J - \lambda L)/(S - \lambda L)$ approaches 1 as $T \rightarrow \infty$, but (3) provides an approximation for the ratio for finite T . This ratio estimate as a function of L and T is displayed in Table 1. We can also use (1) and (2) to develop an approximate expansion directly for the ratio J/S in powers of $1/\sqrt{\lambda L}$. In particular, by (1) and (2),

$$\frac{J}{S} \approx \frac{1 + \sqrt{\frac{c^2}{\lambda L}} \Phi_J}{1 + \sqrt{\frac{c^2}{\lambda L}} \Phi_S} \approx 1 - \sqrt{\frac{c^2}{\lambda L}} (\Phi_S - \Phi_J) + \frac{c^2}{\lambda L} \Phi_S (\Phi_S - \Phi_J).$$

However, we will focus on (3).

The maximum window contents S and J are of course random variables. We also derive rough estimates for the standard deviations, namely,

$$SDV(S) \approx SDV(J) \approx 1.28 \sqrt{\frac{\lambda c^2 L}{2 \log(T/L)}}; \tag{4}$$

see Section 3. Moreover, we would approximate the distributions of S and J , after appropriate normalization, by the type- I extreme-value distribution, i.e., the cdf (cumulative distribution function) $\exp(-e^{-x})$, $-\infty < x < \infty$; see p. 4 of Leadbetter, Lindgren and Rootzén [11] and (10) – (13) below.

We can combine (1), (2), and (4) to obtain estimates of the coefficients of variation (CV , standard deviation divided by the mean) of $S - \lambda L$ and $J - \lambda L$. Keeping only the dominant $\sqrt{2\log(T/L)}$ term in Φ_S and Φ_J , we obtain

$$CV(S - \lambda L) \approx CV(J - \lambda L) \approx \frac{0.64}{\log(T/L)}. \quad (5)$$

Hence, if $T/L = 10^k$, then $CV(S - \lambda L) \approx 0.278/k$. If $k \geq 3$, then the standard deviation of $S - \lambda L$ should be less than 10% of the mean.

In §5 we briefly discuss *discriminating windows* with $L = O(\log T)$, for which the asymptotics depend upon *more* than two parameters λ and c^2 , and thus have more potential for providing information about the input process.

In §6 we make numerical comparisons using computer simulations. In particular, we consider 10 replications each of several renewal processes with $\lambda = 1$, $T = 10^6$ and various values of L . These simulation experiments provide additional support for approximations (1) – (5). Indeed, the last term in Φ_S in (1) was suggested by the simulations. In §7 we state our conclusions.

We were motivated to look for approximations (1) and (2), because we want to compare the open-loop window flow-control schemes to other open-loop flow-control schemes. In particular, in [1] we compare the sliding window to a leaky-bucket-based flow-control scheme. There we conclude that the sliding window admits larger bursts than the leaky bucket for given peak rate and given sustainable rate. To draw this conclusion, we carry out a specific construction: We generate a sample path of a stationary point process I and specify a window length L . The reciprocal of the observed minimum distance between consecutive points is the realized peak rate, and the ratio S/L is the maximum possible sustainable rate, when $S \equiv S(L, T, I)$ is the maximum sliding window content. With this peak rate and sustainable rate, the observed sample path just passes through the sliding window control. We then let the leaky bucket have drain rate r equal to the sustainable rate for the sliding window, and let the bucket capacity be the smallest level such that the observed stochastic sample path just passes through the leaky bucket. With this construction, both controls have the same peak and sustainable rates, and the given sample path just passes through both controls. Finally, with the control parameters so determined, we ask which control scheme allows larger bursts, where a burst is defined as the maximum number of consecutive arrivals at the peak rate. We find that the sliding window allows larger bursts, and we use (1) to estimate the difference. See [1] for more details. Reibman and Berger [14] reach the same conclusions for sample video conferencing sequences.

2. The Central Limit Theorem and Strong Approximations

First, as T gets large, we can apply extreme value theory, as in Leadbetter, Lindgren and Rootzén [11]. In general, though, the limiting behavior depends quite delicately upon the input process I . However, if we can assume that L also is suitably large, then we can hope for significant simplification. As L gets large, in great generality, each increment $I(t+L) - I(t)$ becomes approximately normally distributed by the central limit theorem (CLT), i.e.,

$$\frac{I(t+L) - I(t) - \lambda L}{\sqrt{\lambda c^2 L}} \Rightarrow N(0, 1) \text{ as } L \rightarrow \infty \quad (6)$$

for all t , where \Rightarrow denotes convergence in distribution, $N(0,1)$ denotes a standard (zero mean, unit variance) normal random variable, λ is a positive constant called the *input rate* and c^2 is a positive constant called the *asymptotic variability parameter*.

To have the CLT (6), we need to have the second moment $EI(t)^2$ finite and we need to have the dependence between the increments $I(t+h) - I(t)$ and $I(s) - I(s-h)$ for $s < t$ and $h > 0$ decrease suitably as $t - s$ increases. We assume that we do *not* have infinite second moments or the long-range dependence that would invalidate (6).

If I is a stationary stochastic point process (i.e., has stationary increments) then c^2 is typically the limiting value of the *index of dispersion for counts* (IDC), i.e.,

$$c^2 = I_c(\infty) \equiv \lim_{t \rightarrow \infty} \frac{\text{Var } I(t)}{EI(t)}; \tag{7}$$

see p. 71 of Cox and Lewis [4]. (The role of the second moment condition is clear in (7).) If I is a renewal counting process with i.i.d interrenewal times X_k , then c^2 is the *squared coefficient of variation* of (SCV) of X_1 , i.e.,

$$c^2 = \text{SCV}(X_1) = \frac{\text{Var } X_1}{(EX_1)^2} \tag{8}$$

(It is important that a square appears in the denominator of (8), but not in (7).)

In fact, we need more than (6), and we often have it. In great generality, a properly normalized input process such as I can be approximated by a Brownian motion. Indeed, our key assumption will be that I can be *approximated strongly* by a Brownian motion, i.e., there exists a standard (zero drift, unit variance) Brownian motion (BM) $B \equiv \{B(t) : t \geq 0\}$ such that

$$\sup \left\{ \left| \frac{I(t) - \lambda t}{\sqrt{\lambda c^2}} - B(t) \right| : 0 \leq t \leq T \right\} = O(\log T) \text{ w.p.1 as } T \rightarrow \infty; \tag{9}$$

see Csörgő and Révész [6] and Csörgő, Horváth and Steinebach [5]. These references show that (9) holds in great generality, so that this seems to be a reasonable starting point for us. Note that (9) can be regarded as a form of *functional central limit theorem* (FCLT) generalizing the CLT in (6). To convert (9) into the FCLT, replace t in (9) by nt , T by n , and then divide by \sqrt{n} . This yields

$$\sup \left\{ \left| \frac{I(nt) - \lambda nt}{\sqrt{\lambda c^2 n}} - \frac{B(nt)}{\sqrt{n}} \right| : 0 \leq t \leq 1 \right\} = O\left(\frac{\log n}{\sqrt{n}}\right) \text{ w.p.1 as } n \rightarrow \infty$$

and note that for each n , $\{B(nt)/\sqrt{n} : t \geq 0\}$ is distributed the same as $\{B(t) : t \geq 0\}$. Hence, from (9) we get

$$\frac{I(nt) - \lambda nt}{\sqrt{\lambda c^2 n}} \Rightarrow B(t) \text{ as } n \rightarrow \infty$$

with the convergence being in the function space $D[0, 1]$; see pp. 11 – 15 and 88 – 95 of [6] for further discussion.

It is also significant that the error in (9) can be regarded as best possible; see [5], [6]. For a renewal process with interrenewal times X_k having mean λ^{-1} and SCV c^2 , (9) holds when $Ee^{sX_1} < \infty$ for some $s > 0$; see Corollary 3.1 on p. 1450 of Csörgő, Horváth and Steinebach [5]. (Other strong approximations hold with bigger errors when only $EX_1^p < \infty$ for some $p > 2$; see

[5], [6].) Theorem 3.1 of [5] also provides a means to treat a large class of counting processes with dependent intervals. Moreover, $I(t)$ need not be a counting process; e.g., it could be the input in a fluid model.

In applications, e.g., to communications networks, we are often interested in large L and large T . A key issue in the asymptotics is the way these two parameters grow. We consider this issue in §4.

3. Extreme Values for Brownian Motion

In this section we simply suppose that I can be approximated by a BM . Consistent with (9), we assume that $I(t) = \lambda t + \sqrt{\lambda c^2} B(t)$, where B is a standard BM . In this case, we can apply classical results to describe J and S as $T \rightarrow \infty$ for any fixed L , omitting I from the notation. In particular, by the extreme-value theorem for i.i.d. normal random variables in Theorem 1.5.3 on p. 14 of [11],

$$a(L, T)(J(L, T) - \lambda L - b(L, T)) \Rightarrow Z \text{ as } T \rightarrow \infty, \quad (10)$$

where Z is the classical type- I extreme-value distribution with cdf

$$P(Z \leq x) = \exp(-e^{-x}), \quad -\infty < x < \infty, \quad (11)$$

$$a(L, T) = \sqrt{2 \log(T/L) / \lambda c^2 L} \quad (12)$$

and

$$b(L, T) = \sqrt{\lambda c^2 L} \left[\sqrt{2 \log(T/L)} - \frac{[\log \log(T/L) + \log 4\pi]}{2\sqrt{2 \log(T/L)}} \right]. \quad (13)$$

Note that $EZ = 0.5772$ (Euler's constant), $\text{Var}Z = \pi^2/6 = 1.645$, $SCV(Z) = 4.94$, $\text{median}(Z) = 0.3667$ and $\text{mod}(Z) = 0.9624$ for the type- I extreme-value random variable Z in (10); see Chapter 21 of Johnson and Kotz [8].

We use (10) to develop approximation (2) for the mean of J and thus J itself. In particular, we let

$$J(L, T) \approx \lambda L + b(L, T) + \frac{EZ}{a(L, T)}, \quad (14)$$

which reduces to (2). We also use (10) to obtain an estimate of the variability of J . Based on (10), we have the approximation

$$SDV(J) \approx \frac{SDV(Z)}{a(L, T)}, \quad (15)$$

which yields (4).

We point out that it is not obvious that approximations (14) and (15) ((2) and (4)) will be good. First, we need to approximate I by Brownian motion and, second, we need the extreme-value asymptotics for Brownian motion to perform suitably well for realistic values of T . It is well known that the CLT (6) often yields remarkably good approximations for moderate values of the limiting variable (when there is weak dependence). In contrast, the extreme-value limit

theorems do not yield such good approximations. In part, this is due to the rate of convergence in the extreme-value theorems being remarkably slow; see Section 2.4 of Leadbetter, Lindgren and Rootzén [11]. For example, for the limit (10) the rate of convergence of the normalized cdf's is slower than $1/\log(T/L)$; see (2.4.8) and (2.4.10) of [11]. Hence, even if $T = 10^6$ and $L = 10^3$, good approximations based on (10) are not automatic. Thus, we should be interested in the computer simulations in §6.

As a corollary to (10), we can divide by $a(L,T)$ and obtain the simpler result

$$J(L,T) - \lambda L - \sqrt{2\lambda c^2 L \log(T/L)} \Rightarrow 0 \text{ as } T \rightarrow \infty. \quad (16)$$

When we divide by $\sqrt{\log(T/L)}$ in (16), we obtain

$$\frac{J(L,T) - \lambda L}{\sqrt{\log(T/L)}} \Rightarrow \sqrt{2\lambda c^2 L} \text{ as } T \rightarrow \infty \quad (17)$$

for fixed L . In (16) and (17) convergence in distribution is equivalent to convergence in probability since the limits are deterministic. Note that (16) and (17) yield cruder approximations than (10).

An analog of (17) also exists for S ; see Corollary 1.2.3 on p. 31 of [6] (also using the observation on the bottom of p. 29 and Remark 1.2.1 on p. 30 of [6]). In particular,

$$\frac{S(L,T) - \lambda L}{\sqrt{\log(T/L)}} \rightarrow \sqrt{2\lambda c^2 L} \text{ w.p. 1 as } T \rightarrow \infty. \quad (18)$$

The limits (17) and (18) imply that

$$\frac{S(L,T) - \lambda L}{J(L,T) - \lambda L} \Rightarrow 1 \text{ as } T \rightarrow \infty. \quad (19)$$

Hence, using the Brownian approximation, by (19) S and J are asymptotically equivalent as $T \rightarrow \infty$ for any fixed L . Moreover, (17) and (18) are consistent with S and J being asymptotically approximated by (1) and (2), but we need additional refinement to get the approximations in (1) and (2).

4. The Role of L in the Brownian Motion Approximation

The asymptotics in the previous section apply if we do indeed have a BM approximation. However, the quality of the Brownian approximation for the increments of another process I depends on L being large. Indeed, for the CLT in (6), we required that $L \rightarrow \infty$.

Moreover, in order to have the Brownian extreme-value theory apply to the process I , it is evident that L must go to infinity as T goes to infinity. In particular, we need $L_T/\log T \rightarrow \infty$, where the notation L_T indicates that L depends on T . When $L_T = O(\log T)$, the asymptotics for I depend on more than two parameters λ and c^2 . The case in which $I(t)$ is a partial sum of i.i.d. random variables is discussed in Sections 2.4 and 3.1 of [6]; also see §5 below.

When we stipulate that L_T grows suitably fast as $T \rightarrow \infty$, we can obtain a positive result. The following parallels Theorem 3.1.1 on p. 115 of [6].

Theorem 1: *If I admits the strong approximation (9) and if $L_T \rightarrow \infty$ as $T \rightarrow \infty$ with $0 < L_T < T$, T/L_T nondecreasing, $\log(T/L_T)/\log \log T \rightarrow \infty$ and $L_T/\log T \rightarrow \infty$, then*

$$\frac{S(L_T, T, I) - \lambda L_T}{\beta(L_T, T) \sqrt{\lambda c^2}} \rightarrow 1 \text{ w.p. 1 as } T \rightarrow \infty, \quad (20)$$

where

$$\beta(L_T, T) = (2L_T[\log(T/L_T) + \log \log T])^{1/2}. \tag{21}$$

Proof: Combine Theorem 1.2.1 on p. 30 of [6] with (9), paralleling the proof of Theorem 3.1.1 on p.115 of [6]. It suffices to work with the Brownian motion because our assumption that $L_T/\log T \rightarrow \infty$ as $T \rightarrow \infty$ implies that $(\log T)/\beta(L_T, T) \rightarrow 0$ as $T \rightarrow \infty$. \square

Note that Theorem 1 remains unchanged if we replace $\beta(L_T, T)$ in (21) by $(2L_T \log(T/L_T))^{1/2}$, because of our assumption that $\log(T/L_T)/\log \log T \rightarrow \infty$ as $T \rightarrow \infty$. However, $\beta(L_T, T)$ in (21) arises naturally in the proof. Moreover, a partial result holds even when $\log(T/L_T)/\log \log T \rightarrow r$, $0 < r < \infty$; see Remark 1.2.2 on page 35 of [6]. Hence, there is a basis for (21) in its current form.

Note that, except for the last term, the approximations for S in (1) comes directly from (20); i.e., we use the approximation $S \approx \lambda L_T + \beta(L_T, T)\sqrt{\lambda c^2}$. The last term in (1) is added based on our numerical experience with simulations; see Section 6. It evidently remains to establish a theorem of the form (10) for the sliding window applied to Brownian motion.

To have a more concrete form for approximations, we suppose that $L_T = xT^p$ for fixed x and $0 < p < 1$. Note that the conditions on L_T in Theorem 1 are satisfied if $L_T = xT^p$ for any p , $0 < p < 1$, but *not* for $p = 1$, no matter how small x is.

Corollary: *If I admits the strong approximation (9) and if $L_T = xT^p$, for $x > 0$ and $0 < p < 1$, then*

$$\frac{S(xT^p, T, I) - \lambda xT^p}{\gamma(x, T)\sqrt{\lambda c^2}} \rightarrow 1 \text{ w.p. 1 as } T \rightarrow \infty,$$

where

$$\gamma(x, T) = \left(2xT^p[(1-p)\log T - \log x + \log \log T]\right)^{1/2}.$$

We now obtain a related result for the jumping window J and relate the asymptotics for J and S . In particular, we show that they are asymptotically equivalent in this asymptotic regime and provide support for (2).

Theorem 2: *Under the assumptions of Theorem 1,*

$$\frac{J(L_T, T, I) - \lambda L_T}{b(L_T, T)} \Rightarrow 1 \text{ w.p. 1 as } T \rightarrow \infty \tag{22}$$

for $b(L, T)$ in (13). *If, in addition, $L_T/(\log T)^3 \rightarrow \infty$ as $T \rightarrow \infty$, then*

$$a(L_T, T)[J(L_T, T, I) - \lambda L_T - b(L_T, T)] \Rightarrow Z \text{ as } T \rightarrow \infty \tag{23}$$

for $a(L, T)$, $b(L, T)$ and Z in (11) – (13).

Proof: Combine the strong approximation (9) with the extreme-value result (10). The limit (10) holds as $T \rightarrow \infty$ for fixed L , but we need a limit as $T \rightarrow \infty$ and $L \rightarrow \infty$. By the properties of Brownian motion, (10) holds uniformly in L provided that $T/L \rightarrow \infty$; i.e., given that T is an integer multiple of L , $[J(L, T) - \lambda L]/\sqrt{L}$ for Brownian motion with drift λ and variance coefficient 1 is distributed the same as $J(1, T/L)$ for Brownian motion with drift 0 and variance coefficient 1. For the case in which I in $J(L_T, T, I)$ is Brownian motion, (22) follows from (10) by dividing by $a(L_T, T)b(L_T, T)$. For the case of a general I in $J(L_T, T, I)$, we apply (9). Under the assumptions of Theorem 1, $L_T/\log T \rightarrow \infty$, so that $(\log T)/b(L_T, T) \rightarrow 0$ as $T \rightarrow \infty$. For (23),

we use the additional assumption to obtain $a(L_T, T) \log T \rightarrow 0$ as $T \rightarrow \infty$. \square

In Theorems 1 and 2 we have provided asymptotics that can serve as the basis for approximations. If better numerical results are desired, especially for smaller values of T and L , then the more involved approximation method of Naus [12], [13] is useful.

5. Discriminating Windows

The appeal of the asymptotics in §4 is the simple formulas depending on the process I only via the two parameters λ and c^2 . However in some circumstances we may want to learn more about the process I from window processes. For this purpose, it is natural to be guided by §2.4 of [6], and let

$$L_T = x \log T. \quad (24)$$

For the case of partial sums of i.i.d. random variables, i.e., $I(t) = X_1 + \dots + X_{[t]}$, where $[t]$ is the integer part of t ,

$$\frac{S(x \log T, T, I)}{x \log T} \rightarrow \alpha(x) \text{ w.p. 1 as } T \rightarrow \infty, \quad (25)$$

where

$$\alpha(x) = \sup\{x: \inf_t e^{-tx} E e^{tX_1} \geq e^{-1/x}\}; \quad (26)$$

see Theorem 2.4.3 on p. 98 of [6]. Moreover, by Theorem 2.4.5 of [6], the function $\alpha(x)$ in (26) uniquely determines the distribution function of X_1 . More generally, the limit (25) regarded as a function of x seems promising for characterizing the entire process I . This seems to be a promising direction for research.

6. Numerical Comparisons

In this section we use computer simulation to investigate the approximations in (1) – (5). In all cases we let $\lambda = 1$. We start by letting $T = 10^6$.

We first consider the Poisson process, which has $c^2 = 1$. Tables 2 and 3 display results for S and J based on two experiments, each with 10 independent replications. We let L range from 20 to 5000. Since $T = 10^6$ and $\log T = 13.8$, in this range for L we clearly have $L \ll T$, but we only have $\log T \ll L$ for larger values of L , say $L \geq 100$.

From the first experiment, we saw that the approximation for S provided by Theorem 1 performs reasonably well, but that it can be improved by subtracting the standard deviation in (4). This yields the approximation for S or ES in (1). This refinement was found to help in other experiments as well, so we include it in (1). However, it still remains to develop theoretical justification for this refinement. Note that the standard deviation in (4) is of the same order as $1/a(L, T)$ in (10) – (14), so that it is natural to anticipate a refinement for S of this form. To confirm this refinement, we conducted the second experiment with different random seeds.

In Table 2 we display the error in the mean of S for each simulation experiment. This is the approximation value (1) minus the simulation estimate. Notice that the approximation for the mean performs well for the entire range of window lengths. Also notice that the estimated errors in the mean are quite a bit smaller than the standard deviation when L is not too small (e.g., for

$L \geq 100$), both predicted by (4) and estimated by simulation. Hence, in the region $\log T \ll L \ll T$, the approximation is accurate to within the noise, e.g., the maximum of the strong approximation error $O(\log T) = O(13.8)$ and one standard deviation.

Table 3 describes similar results for the jumping window, based on the same simulation experiments. Again, the error in the mean is the approximate value minus the simulation estimate. Here no additional refinement was deemed necessary, so approximation (2) for J follows directly from Theorem 2 and (10) – (14). As with the sliding window, the approximation for the mean performs quite well for a Poisson process over the entire range of window lengths. Again the error in the mean is less than one standard deviation for $L \geq 100$.

From Tables 2 and 3, we see that there is considerable variability in the standard deviation estimates. However, (4) is evidently a reasonable estimate of the standard deviation of both S and J .

To test the invariance principle, we also performed the identical simulation experiment for two rate-1 renewal processes with common SCV, $c^2 = 15.4$. One renewal process is an on-off process with a deterministic distance of 0.1 between arrivals during an on period. The on period is geometrically distributed with mean 1, so that the mean number of arrivals during an on period is 10. As usual, the off period is exponentially distributed. An interarrival time is the mixture of 0.1 and 0.1 plus the exponential off period. The mean of the exponential random variable is chosen so that the mean interarrival time is 1.0. This makes the SCV $c^2 = 15.4$.

The other renewal process has interarrival times equal to a constant 0.1 plus a hyperexponential (H_2) random variable, which has a distribution which is a mixture of two exponential distributions; i.e., the H_2 distribution has density

$$f(t) = p\lambda_1 e^{-\lambda_1 t} + (1-p)\lambda_2 e^{-\lambda_2 t}, \quad t \geq 0. \quad (27)$$

The H_2 distribution is given balanced means, i.e., $p\lambda_1^{-1} = (1-p)\lambda_2^{-1}$. The remaining two parameters are chosen so that the overall distribution has mean 1 and SCV $c^2 = 15.4$.

Tables 4 and 5 display the results for these two renewal processes with common SCV $c^2 = 15.4$. Table 4 displays results for the sliding window, while Table 5 displays results for the jumping window. As in Tables 2 and 3, the simulation estimates are based on ten independent replications for each renewal process. A second experiment of ten independent replications yielded similar results.

From Tables 4 and 5, we see that there are significant differences between these two renewal process. However, for larger values of L , these differences are relatively small compared to the standard deviation. For $L \leq 500$, though, the differences are dramatic. For larger values of L , say $L \geq 500$, the results are roughly consistent with the invariance principle. The approximate mean overestimates the $(D + H_2)$ mean values by about the same amount it underestimates the on-off mean values. The error in the mean relative to the standard deviation decreases significantly as L increases. At about $L \geq 700$, the errors are no more than one standard deviation.

For the smaller values of L , we see that the invariance principle is *not* appropriate. Consistent with §5, we see that for $L \leq 100$ the two renewal processes behave very differently. Both the means and the standard deviations differ by a factor of two to five in this range. It is interesting that the ratio of the standard deviation estimates of the $D + H_2$ renewal process to the on-off renewal process increases from 0.22 at $L = 20$ to 1.8 for $L = 5000$. As with the mean, the approximation falls nicely between the two standard deviation estimates.

We also performed similar experiments with less bursty variants of these same two renewal processes. We found that the approximations performed better when the SCV is decreased from

15.4 to 4.0. As might be anticipated, this case falls in between the cases $c^2 = 1$ and $c^2 = 15.4$ that we have described in detail.

Next, Table 6 compares the approximation for the ratio $(J - \lambda L)/(S - \lambda L)$ in (3) with simulation estimates for the three renewal processes considered above. Note that the approximation is the same for all three processes. The estimates are the sample means and sample standard deviations of the observed ratio. The average errors over the 14 values with $L \geq 500$ are -0.015 for the Poisson process, -0.001 for the on-off renewal process and -0.027 for the $(D + H_2)$ -renewal process. Notice that the estimated standard deviations are about 0.09 for the larger L values and about 0.05 for the smaller L values. The estimated errors in the mean are quite small compared to these standard deviation estimates.

Finally, it is interesting to see how the approximations perform for smaller values of T than 10^6 . To give an idea, Tables 7 and 8 display results for the sliding and jumping windows, respectively, applied to the Poisson process for $T = 10^3$ and $T = 10^4$. Again, the simulation estimates are based on ten independent replications and the error in the mean is the approximation minus the sample mean. Consistent with Tables 2 and 3, the approximation for the means tend to be slightly low for small values of L . Table 7 shows that for $T = 10^4$ and $L \geq 1000$, where $T/L < 10$, the approximation degrades, but it still is within one standard deviation. Table 8 shows that the approximation for the jumping window degrades more for large L . The ratio estimates are also highly variable in this region. For example, for $T = 10^4$ and $L = 2.5 \times 10^3$, the sample mean and variance of $(J - \lambda L)/(S - \lambda L)$ were 1.19 and 1.83, respectively. This should not be surprising, since $T/L = 4$ in this case. Contrary to the conditions of Theorems 1 and 2, $\log(T/L)$ is *not* large compared to $\log \log T$ in this range, so this degradation in performance can be anticipated.

However, for L neither very small nor very large, Tables 7 and 8 show that the approximations still perform well.

7. Conclusions

Our purpose in this paper was to develop simple approximations for the maximum sliding and jumping window content associated with a general input process I . Since our intended application is high-speed communication networks, we are interested in relatively long sample paths (very large values of T) and large window length L , with T much larger than L . Hence, it is natural to consider applications based on asymptotics.

We proposed approximating the general input process I by a Brownian motion. This approximation depends on the general input process I through only two parameters: its rate λ and the limiting value of its index of dispersion, c^2 in (7). Strong approximations allow us to do this for the entire process yielding an error of $O(\log T)$; see (9).

We then consider extreme-value asymptotics for Brownian motion in order to obtain the basic approximations in (1) and (2). These approximations were supported by Theorems 1 and 2, which established limits as both $T \rightarrow \infty$ and $L \rightarrow \infty$ with $L/\log T \rightarrow \infty$.

We developed the final form of the approximations for ES in (1) and provided additional support for the other approximations by making numerical comparisons with computer simulations. (However, a refinement of Theorem 1 supporting approximation (1) is still needed.) When $T = 10^6$ and $L = 10^3$, the basic approximations (1)–(4) perform well. However, when $T = 10^6$ and $L = 50$, and L is no longer much bigger than $\log T$, the simulations confirm that the invariance principle can break down, as predicted by the theory (see §5). This break down was illustrated by the example of the two renewal processes with SCV 15.4.

Overall, the approximations seem remarkably effective for the parameter range indicated by the theory. Once again, we see that statistical regularity can be captured by appropriate probability limit theorems.

References

- [1] Berger, A.W. and Whitt, W., Comparison of two traffic descriptors for ATM connections in a broadband ISDN, (in preparation).
- [2] Bertsekas, D. and Gallager, R., *Data Networks*, Prentice Hall, Englewood Cliffs, NJ, (1987).
- [3] Budka, K.C., Stochastic monotonicity and concavity properties of rate-based flow control mechanisms, *IEEE Trans. Aut. Control*, **38**, to appear (1993).
- [4] Cox, D.R. and Lewis P.A.W., *The Statistical Analysis of Series of Events*, Methuen, London, (1966).
- [5] Csörgő, M, Horváth, L., and Steinebach, J., Invariance principles for renewal processes, *Ann. Probab*, **15**, 1441-1460, (1987).
- [6] Csörgő, M. and Révész, P., *Strong Approximations in Probability and Statistics*, Academic Press, New York, (1981).
- [7] Erdős, P. and Rényi, A., On a new law of large numbers, *J. Analyse Math.* **23**, 123-131, (1970).
- [8] Johnson, N.L. and Kotz, S., *Distributions in Statistics: Continuous Univariate Distributions-1*, Wiley, New York, (1970).
- [9] Komlós, J., Major, P., and Tusnády, G., An approximation of partial sums of independent rv's and the sample D.F-I., *Z. Wahrscheinlichkeitsth. Verw, Geb.* **32**, 111-131, (1975).
- [10] Komlós, J., Major, P., and Tusnády, G., An approximation of partial sums of independent random variables and the sample D.F-II., *Z. Wahrscheinlichkeitsth. Verw, Geb.* **34**, 33-58, (1976).
- [11] Leadbetter, M.R., Lindgren, G., and Rootzén, H., *Extremes and related properties of random sequences and processes*, Springer-Verlag, New York (1983).
- [12] Naus, J.I., Approximations for distributions of scan statistics, *J. Amer. Stat. Assoc.* **77**, 177-183, (1982).
- [13] Naus, J.I., Scan statistics, in *Encyclopedia of Statistical Sciences*, S. Kotz and N.L. Johnson (eds), **8**, 281-284, Wiley, New York, (1988).
- [14] Reibman, A.R. and Berger, A.W., Traffic descriptors for VBR video conferencing over ATM networks, *IEEE/ACM Trans. on Networking*, to appear.
- [15] Roberts, J., (ed.) *Performance Evaluation and Design of Multiservice Networks*, COST 224 Final Report, Commission of the European Communities, Luxembourg, (1992).

L	T	10^5	10^6	10^7	10^8	10^9	10^{10}
50		0.834	0.856	0.872	0.885	0.894	0.902
100		0.822	0.848	0.866	0.880	0.891	0.900
500		0.787	0.826	0.850	0.868	0.881	0.892
1000		0.767	0.813	0.842	0.862	0.877	0.888
5000		0.700	0.777	0.819	0.846	0.865	0.879
10000		0.656	0.756	0.806	0.837	0.858	0.874

Table 1: The estimated ratio $(J-\lambda L)/(S-\lambda L)$ in (3) as a function of L and T .

window length L	approximations		simulation estimates					
	mean (1)	std. dev. (4)	first experiment			second experiment		
			mean	std. dev.	error in mean	mean	std. dev.	error in mean
20	42	1.2	45	1.8	-3	44	0.9	-2
40	70	1.8	73	1.8	-3	72	1.6	-2
60	96	2.2	99	2.4	-3	98	2.0	-2
80	121	2.6	125	2.1	-4	122	2.2	-1
100	146	3.0	148	1.5	-2	147	1.8	-1
200	262	4.4	262	3.2	0	263	4.5	-1
300	375	5.5	376	3.5	-1	375	4.3	0
400	485	6.5	486	5.9	-1	484	5.0	1
500	594	7.3	594	6.3	0	591	5.6	3
600	702	8.1	703	8.1	-1	697	7.2	5
700	809	8.9	809	8.2	0	802	7.3	7
800	915	9.6	918	11.6	-3	910	10.2	5
900	1021	10.3	1024	8.8	-3	1016	9.6	5
1000	1127	10.9	1128	8.6	-1	1118	6.9	9
1500	1652	13.7	1654	22.3	-2	1645	10.5	7
2000	2172	16.2	2173	23.6	-1	2164	12.7	8
2500	2689	18.5	2692	29.6	-3	2685	21.1	4
3000	3204	20.6	3206	22.5	-2	3199	24.9	5
3500	3718	22.5	3718	22.2	0	3721	26.6	-3
4000	4231	24.4	4223	19.0	8	4228	29.1	3
4500	4743	26.1	4732	15.4	11	4734	38.2	9
5000	5254	27.8	5247	19.7	7	5237	35.9	17

Table 2. A comparison of approximations with simulation estimates for the mean and standard deviation of the maximum *sliding* window content of a rate-1 Poisson process for a total time period $T = 10^6$ as a function of the window length L . The simulation estimates are based on two experiments, each of ten independent replications.

window length L	approximations		simulation estimates					
	mean (2)	std. dev. (4)	first experiment			second experiment		
			mean	std. dev.	error in mean	mean	std. dev.	error in mean
20	39	1.2	42	1.9	-3	41	0.9	-2
40	66	1.8	69	2.9	-3	68	1.6	-2
60	91	2.2	93	2.9	-2	93	2.2	-2
80	115	2.6	117	2.6	-2	117	2.3	-2
100	139	3.0	141	3.5	-2	141	4.3	-2
200	252	4.4	253	4.4	-1	251	4.3	1
300	362	5.5	364	4.6	-2	365	5.0	-3
400	470	6.5	474	9.4	-4	471	4.2	-1
500	577	7.3	579	6.6	-2	575	6.1	2
600	684	8.1	688	7.7	-4	684	6.1	0
700	789	8.9	793	8.9	-4	785	6.9	4
800	894	9.6	896	8.9	-2	887	6.8	7
900	999	10.3	1001	10.4	-2	999	8.4	0
1000	1103	10.9	1100	12.0	3	1097	9.0	6
1500	1622	13.7	1624	15.7	-2	1627	14.2	-5
2000	2137	16.2	2134	19.4	3	2133	17.2	4
2500	2650	18.5	2661	31.3	-11	2654	12.7	-4
3000	3161	20.6	3164	25.8	-3	3167	27.3	-6
3500	3672	22.5	3669	23.9	3	3663	21.2	9
4000	4181	24.4	4168	19.0	13	4181	17.0	0
4500	4689	26.1	4696	24.7	-7	4684	31.7	5
5000	5197	27.8	5181	26.1	16	5178	18.1	19

Table 3. A comparison of approximations with simulation estimate for the mean and standard deviation of the maximum *jumping* window content of a rate-1 Poisson process for a total time period $T = 10^6$ as a function of the window length L . The simulation estimates are based on two experiments, each of ten independent replications.

window length L	approximations		simulation estimates					
			mean		error in mean		std. dev.	
	mean (1)	std. dev. (4)	on-off	$D+H_2$	on-off	$D+H_2$	on-off	$D+H_2$
20	106	4.8	152	58	-46	48	7.3	1.6
40	158	7.1	202	101	-44	57	7.7	3.0
60	202	8.8	251	140	-49	62	9.6	2.9
80	242	10.3	287	179	-45	63	21.3	2.9
100	279	11.7	330	214	-51	65	19.0	3.7
200	445	17.2	483	380	-38	65	10.5	7.0
300	593	21.6	645	531	-52	62	36.3	19.2
400	733	25.4	774	674	-41	59	31.4	20.1
500	868	28.8	919	816	-51	52	38.7	16.4
600	999	31.9	1035	951	-36	48	40.6	14.5
700	1127	34.9	1157	1078	-30	49	42.2	16.5
800	1252	37.6	1287	1197	-35	55	35.4	19.0
900	1376	40.2	1418	1327	-42	49	43.0	26.9
1000	1499	42.7	1542	1454	-43	45	51.6	27.4
1500	2095	53.9	2134	2037	-39	58	53.9	29.6
2000	2674	63.7	2691	2650	-17	24	55.7	34.7
2500	3242	72.5	3275	3209	-33	33	57.9	75.3
3000	3802	80.7	3829	3799	-27	3	67.1	92.9
3500	4356	88.3	4398	4332	-42	24	48.4	110.0
4000	4906	95.6	4943	4851	-37	55	88.4	135.3
4500	5452	102.5	5471	5410	-19	42	92.5	144.2
5000	5995	109.1	6008	5966	-13	29	80.9	145.4

Table 4. A comparison of approximations with simulation estimates for the mean and standard deviation of the maximum *sliding* window content of two rate-1 renewal processes with SCV $c^2=15.4$ for a total time period of length $T = 10^6$ as a function of the window length L . The simulation estimates are based on ten independent replications.

window length L	approximations		simulation estimates					
			mean		error in mean		std. dev.	
	mean (2)	std. dev. (4)	on-off	$D+H_2$	on-off	$D+H_2$	on-off	$D+H_2$
20	95	4.8	135	55	-40	40	8.2	2.2
40	141	7.1	186	96	-45	45	8.8	2.8
60	181	8.8	227	135	-46	46	11.4	3.0
80	218	10.3	257	170	-39	48	13.0	2.9
100	252	11.7	306	203	-54	49	21.9	4.3
200	405	17.2	448	362	-43	43	22.8	10.1
300	544	21.6	591	507	-47	37	24.4	15.2
400	676	25.4	728	647	-52	29	33.5	19.2
500	804	28.8	847	771	-43	33	22.9	15.6
600	928	31.9	958	897	-30	31	36.3	27.9
700	1050	34.9	1077	1022	-27	28	29.7	25.0
800	1170	37.6	1206	1151	-36	19	46.1	29.0
900	1289	40.2	1318	1263	-29	26	39.1	24.2
1000	1406	42.7	1434	1388	-28	18	34.4	32.5
1500	1980	53.9	1991	1954	-11	26	59.2	20.6
2000	2539	63.7	2566	2518	-27	21	65.7	82.1
2500	3089	72.5	3091	3070	-2	19	49.0	76.0
3000	3634	80.7	3644	3665	-10	-31	58.2	116.5
3500	4173	88.3	4170	4138	3	35	60.6	80.6
4000	4709	95.6	4761	4673	-52	36	61.2	103.0
4500	5243	102.5	5283	5228	-40	15	59.6	142.0
5000	5773	109.1	5801	5774	-28	-1	117.7	109.4

Table 5. A comparison of approximations with simulation estimates for the mean and standard deviation of the maximum *jumping* window content of two rate-1 renewal processes with $SCV\ c^2=15.4$ for a total time period of length $T = 10^6$ as a function of the window length L . The simulation estimates are based on ten independent replications.

window length L	app- rox. ratio (3)	simulation estimates								
		Poisson			on-off, $c^2=15.4$			$D+H_2$, $c^2=15.4$		
		m e a n	std. dev.	e r r o r	m e a n	std dev	e r r o r	m e a n	std. dev.	e r r o r
20	.865	.856	.050	.009	.878	.073	-.013	.922	.050	-.057
40	.858	.871	.045	-.013	.904	.076	-.046	.928	.042	-.070
60	.854	.871	.058	-.017	.878	.067	-.024	.936	.046	-.082
80	.851	.883	.050	-.032	.862	.096	-.011	.912	.033	-.061
100	.848	.885	.075	-.037	.900	.071	-.052	.897	.030	-.049
200	.839	.816	.086	.023	.876	.075	-.037	.903	.028	-.064
300	.833	.870	.055	-.037	.848	.084	-.015	.899	.040	-.066
400	.829	.845	.060	-.016	.878	.062	-.049	.902	.054	-.073
500	.826	.826	.080	.000	.831	.056	-.005	.859	.041	-.033
600	.823	.871	.089	-.048	.827	.093	-.004	.844	.070	-.021
700	.820	.832	.093	-.012	.828	.049	-.008	.851	.069	-.031
800	.818	.798	.079	.020	.833	.067	-.015	.884	.068	-.066
900	.815	.862	.084	-.047	.812	.094	.003	.854	.088	-.039
1000	.813	.827	.088	-.014	.804	.074	.009	.856	.085	-.043
1500	.806	.881	.085	-.075	.775	.080	.031	.848	.066	-.042
2000	.799	.814	.074	-.015	.822	.095	-.023	.795	.103	.004
2500	.794	.843	.108	-.049	.765	.079	.029	.802	.040	-.008
3000	.790	.841	.090	-.051	.781	.100	.009	.831	.090	-.041
3500	.786	.742	.070	.044	.748	.079	.038	.772	.094	.014
4000	.783	.803	.082	-.020	.811	.085	-.028	.794	.072	-.011
4500	.780	.791	.102	-.011	.813	.094	-.033	.801	.091	-.021
5000	.777	.760	.104	.017	.794	.090	-.017	.811	.122	-.034
average errors for $L \geq 500$				-.015		-.001				-.027

Table 6. A comparison of the approximation for the normalized ratio $(J-\lambda L)/(S-\lambda L)$ with simulation estimates for the Poisson process, the on-off renewal process, and the $D+H_2$ renewal process for a time period $T = 10^6$ as a function of the window length L . The simulation estimates are based on ten independent replications. The error is the approximation minus the estimated mean.

window length L	$T = 10^3$				$T = 10^4$			
	approx.	simulation estimates			approx.	simulation estimates		
	mean (1)	mean	error in mean	std. dev.	mean (1)	mean	error in mean	std. dev.
10	20	20.4	-0.4	1.5	22	24.9	-2.9	1.6
20	33	34.3	-1.3	2.1	37	38.3	-1.3	2.3
40	57	57.5	-0.5	2.5	62	62.8	-0.8	2.3
60	80	78.6	1.4	2.6	87	86.6	0.4	5.1
80	102	101.7	0.3	3.3	110	110.6	-0.6	6.9
100	123	122.4	0.6	4.5	133	132.8	0.2	6.9
200	228	228.1	-0.1	9.2	243	241.3	1.7	7.8
300	329	328.1	0.9	11.2	350	347.5	2.5	12.5
400	429	427.1	1.9	13.9	456	452.2	3.8	13.9
500	527	526.4	0.6	17.9	561	558.5	2.5	16.1
1000					1076	1062.2	13.8	22.5
2000					2092	2065.2	26.8	34.9
4000					4099	4054.7	44.3	56.7
5000					5094	5048.2	45.8	64.5

Table 7. A comparison of approximations with simulation estimate for the mean of the maximum *sliding* window content of a rate-1 Poisson process for total time periods $T = 10^3$ and $T = 10^4$ as a function of the window length L . The simulation estimates are based on ten independent replications.

window length L	$T = 10^3$				$T = 10^4$			
	app- rox.	simulation estimates			app- rox.	simulation estimates		
	mean (2)	mean	error in mean	std. dev.	mean (2)	mean	error in mean	std. dev.
10	18	18.8	-0.8	1.2	20	21.4	-1.4	1.6
20	30	30.5	-0.5	2.2	34	34.5	-0.5	1.3
40	53	52.3	0.7	3.3	58	58.1	-0.1	1.9
60	74	73.3	0.7	3.8	81	79.9	1.1	3.6
80	96	94.8	1.2	4.5	104	102.8	1.2	5.1
100	116	113.6	2.4	5.2	126	124.8	1.2	3.8
200	218	216.1	1.9	9.0	233	230.2	2.8	7.8
300	318	314.4	3.6	12.7	337	335.4	1.6	13.4
400	418	411.7	6.3	14.0	441	437.3	3.7	18.2
500	517	504.4	12.6	14.1	543	542.9	0.1	17.3
1000					1052	1042.3	9.7	24.8
2000					2057	2047.6	9.4	33.1
4000					4056	4018.4	37.6	39.2
5000					5053	5013.9	39.1	59.0

Table 8. A comparison of approximations with simulation estimates for the mean of the maximum *jumping* window content of a rate-1 Poisson process for total time periods $T = 10^3$ and $T = 10^4$ as a function of the window length L . The simulation estimates are based on ten independent replications.