Use of the supremum distribution of Gaussian processes in queueing analysis with long-range dependence and self-similarity

Jinwoo Choe
Department of Electrical and Computer Engineering,
University of Toronto, Toronto, Ontario M5S3G4, Canada
E-mail: jinwoo@comm.toronto.edu, Tel. +1 416 978 7001.

Ness B. Shroff
School of Electrical and Computer Engineering,
Purdue University, West Lafayette, IN 47907-1285, U.S.A.
E-mail: shroff@purdue.edu, Tel. +1 765 494 3471.

Abstract In this paper we study the supremum distribution of a general class of Gaussian processes $\{X_t : t \geq 0\}$ with stationary increments. This distribution is directly related to the steady state queue length distribution of a queueing system. Hence, its study is also important for different queueing applications such as delay analysis in communication networks. Our study is based on *Extreme Value Theory* and we show that $\log \mathbb{P}(\{\sup_{t\geq 0} X_t > x\}) + m_x/2$ asymptotically grows at most (on the order of) $\log x$, where m_x is the reciprocal of the maximum (normalized) variance of X_t. This result is considerably stronger than the existing results in the literature based on *Large Deviation Theory*. We further show that this improvement can be critical in characterizing the asymptotic behavior of $\mathbb{P}(\{\sup_{t\geq 0} X_t > x\})$. Our results cover a large class of self-similar, short range dependent, and long-range dependent Gaussian processes.

Keywords steady state queue length distribution, stationary increments, overflow probability, long range dependence, self-similarity, fractal Brownian motion, queueing theory.

1 INTRODUCTION

In this paper, we study the supremum distribution of a Gaussian process having *stationary increments*. In general, a stochastic process $\{X_t : t \geq 0\}$ is said to have stationary increments if the distribution of $X_{t+\tau} - X_t$ depends only on the time difference τ, and not on t. The study of the supremum distribution (the distribution of $\sup_t X_t$) of stochastic processes with stationary increments has attracted considerable interest, in large part because of its direct relation to the steady-state queue length distribution of a queueing system (e.g., see [2, 10, 12, 19]). Consider a queueing system, such as the one shown in Figure 1. Let Λ_t be defined in such a way that $\Lambda_t - \Lambda_s$ is the amount of fluid that arrives into the system during the time interval $(s,t]$. Similarly, we define M_t to be a function of t such that $\mathrm{M}_t - \mathrm{M}_s$ is the maximum amount of fluid that can be served during the time interval $(s,t]$. Then assuming that the queue is empty at $t = 0$, Q_t, the amount of fluid in the system (workload) at time t, can be expressed as $Q_t = \sup_{0 \leq s \leq t} (\mathrm{N}_t - \mathrm{N}_s)$, where $\mathrm{N}_t := \Lambda_t - \mathrm{M}_t$ (e.g., see [10, 12]).

If we assume that Λ_t and M_t are independent stochastic processes with stationary increments, then

$$\begin{aligned} \mathbb{P}(\{Q_t > x\}) &= \mathbb{P}\left(\left\{\sup_{0\leq s\leq t} (\mathrm{N}_t - \mathrm{N}_s) > x\right\}\right) \\ &\to \mathbb{P}\left(\left\{\sup_{s\geq 0} (\mathrm{N}_0 - \mathrm{N}_{-s}) > x\right\}\right), \quad \text{as } t \to \infty. \end{aligned} \tag{1}$$

Hence, $\mathbb{P}(\{Q > x\}) := \lim_{t\to\infty} \mathbb{P}(\{Q_t > x\}) = \mathbb{P}\left(\left\{\sup_{t\geq 0} (\mathrm{N}_0 - \mathrm{N}_{-t}) > x\right\}\right)$. So if we define $X_t := \mathrm{N}_0 - \mathrm{N}_{-t}$, then $\{X_t : t \geq 0\}$ is a stochastic process with stationary increments, and $\mathbb{P}(\{Q > x\}) = \mathbb{P}\left(\left\{\sup_{t\geq 0} X_t > x\right\}\right)$. For notational simplicity, henceforth, we write $\langle w\rangle_\Theta := \sup_{\theta\in\Theta} w_\theta$, where the index set Θ will be omitted when it covers the entire domain on which w_θ is defined.

A large body of the work devoted to the study of the supremum distribution has focused on the asymptotic tail behavior of this distribution; i.e., the asymptotic behavior of $\mathbb{P}(\{\langle X\rangle > x\})$ (or equivalently $\mathbb{P}(\{Q > x\})$). The theory of *Large Deviations* has been widely used, providing very general and elegant results on the asymptotic behavior of $\log \mathbb{P}(\{\langle X\rangle > x\})$ [8, 10]. For example, in [10], using Large Deviation techniques, it has been shown for a

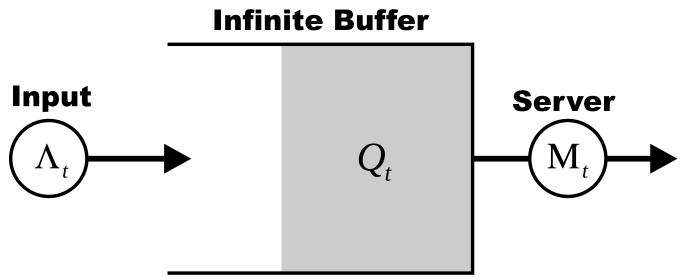

Figure 1: A fluid queueing system with an infinite buffer and a server.

large class of stochastic processes that

$$\log \mathbb{P}(\{\langle X \rangle > x\}) \sim -\eta x, \tag{2}$$

where the asymptotic decay rate, η, is a positive constant that can usually be found in closed form, and $f(x) \sim g(x)$ means that for any $\epsilon > 0$, there exists an x_o such that for all $x > x_o$, $f(x)$ lies in the (closed) interval enclosed by $(1-\epsilon)g(x)$ and $(1+\epsilon)g(x)$. (Note that this is a more general definition of similarity ($\sim$) than the typical definition given by $\lim_{x\to\infty} g(x)/f(x) = 1$. With this new definition, the similarity between two functions $f(x)$ and $g(x)$ can be well defined, even when there are arbitrarily large values of x for which $f(x) = g(x) = 0$. This definition will required for some of the later results in the paper.). However, for many important types of processes, such as *self-similar* or other *long-range dependent* processes [3, 11], the relationship given by (2) may not hold. To address this problem, in [8], the generality of Large Deviation techniques was exploited, and the above result was extended through an elegant scaling technique to obtain

$$\log \mathbb{P}(\{\langle X \rangle > x\}) \sim -q(x), \tag{3}$$

where $q(x)$ is some increasing function of x, which may not be linear in x. However, the great generality of the results based on Large Deviation techniques do come at a cost: poor "resolution." This is because the log-similarity relation given by (3) captures only the leading (most rapidly growing) term of $\log \mathbb{P}(\{\langle X \rangle > x\})$. For example, if $q(x) = x$ satisfies (3), then so does $q(x) = x + \sqrt{x}$, even though it is a very different function of x. Therefore, approximations for $\mathbb{P}(\{\langle X \rangle > x\})$ based on (3) should be used with some caution, since (3) provides relatively weak knowledge of the asymptotic behavior of these approximations. In order to address this difficulty, our *objective* in this paper is to focus on a large class of *Gaussian processes* (including many types of long-range dependent processes) and develop a considerably stronger asymptotic relationship.

Recently, Gaussian processes have received a lot of attention for the modeling and analysis of queueing behavior in high-speed networks [1, 5, 6, 14, 16]. There are many reasons for this. Due to the huge link capacity of high-speed networks, thousands of applications are likely to be served at a network node. This suggests that by appealing to the *Central Limit Theorem*, we can accurately characterize the input process as a Gaussian process. In earlier studies with traffic modeling we have empirically found that, typically, a couple of hundred multiplexed sources are sufficient for the traffic to be modeled as a Gaussian process [5]. Moreover, when a large number of sources are multiplexed, characterizing the input process with traditional Markovian models results in computational infeasibility problems [18] that are not encountered for Gaussian processes. Finally, recent network traffic studies suggest that certain types of network traffic may exhibit self-similar or more generally the asymptotic self-similar type of long-range dependence [3, 11], and various Gaussian processes can be used to model such processes.

We assume that $\{X_t : t \geq 0\}$ is a Gaussian process with stationary increments such that $X_0 = 0$, and define $\kappa := -\mathbb{E}\{X_t\}/t$ and $v(t) := \mathrm{Var}\{X_t\}$. Since X_t is Gaussian, $\mathbb{P}(\{X_t > x\})$ can be expressed in terms of $v(t)$, κ, and the standard Gaussian tail function $\Psi(w) := \int_w^\infty e^{-z^2/2}dz \Big/ \sqrt{2\pi}$, as $\mathbb{P}(\{X_t > x\}) = \Psi\left((x+\kappa t)/\sqrt{v(t)}\right)$. Assuming that $v(t)/t^2 \to 0$ as $t \to \infty$, it is not difficult to see that $v(t)/(x+\kappa t)^2$ should attain its maximum value at some finite $t = t_x$. Therefore, the probability $\mathbb{P}(\{X_t > x\})$ is also maximized at $t = t_x$. The qualitative statement "rare events take place only in the most probable way" (e.g. see [8]) suggests that $\mathbb{P}(\{X_t > x\})$ at $t = t_x$ would be a good lower bound approximation for $\mathbb{P}(\{\langle X\rangle > x\})$, and in fact similar ideas have already been used in different ways to analyze and approximate the tail probability [5, 8, 10]. In [5, 6], we provide a rigorous asymptotic result that theoretically supports the above qualitative statement. This result is both generalized and strengthened by Proposition 3 in this paper. Also, in [5, 6], for a fairly large class of Gaussian processes where the tail probability is asymptotically exponential, i.e., $\mathbb{P}(\{\langle X\rangle > x\}) \sim Ae^{-\eta x}$, we have obtained a tight upper bound on the asymptotic constant A (note that

η in this equation is the same as in (2)). In the course of that study we found that $m_x := (x + \kappa t_x)^2 / v(t_x)$ (the reciprocal of the maximum value of $v(t)/(x + \kappa t)^2$) contains important information about the shape of the tail probability curve, and that $e^{-m_x/2}$ asymptotically bounds the tail probability from above. For the class of Gaussian processes considered in [5, 6], we also found through an extensive simulation study, using importance sampling, that $e^{-m_x/2}$ provides a very accurate estimate of the tail probability over a wide range of queue lengths x, including small values of x.

In this paper, we consider a more general class of Gaussian processes, including a large class of long-range dependent processes, to show that

$$\log \mathbb{P}(\{\langle X \rangle > x\}) + \frac{m_x}{2} \in O(\log x), \tag{4}$$

where $O(f(x))$ denotes the set of functions $g(x)$ such that there exists a $c > 0$ that satisfies $|g(x)| \leq c|f(x)|$ for all sufficiently large x. Observe that (4) characterizes $\log \mathbb{P}(\{\langle X \rangle > x\})$ in much more detail than (3). Further, (4) suggests that the asymptotic behavior of $\log \mathbb{P}(\{\langle X \rangle > x\})$ is very similar to that of $-m_x/2$, and that the difference between them is asymptotically either a constant (as found in [5, 6] in a more restrictive setting) or a very slowly growing function of x. Therefore, (4) provides more information on the asymptotic behavior of $\mathbb{P}(\{\langle X \rangle > x\})$ than (3), and suggests that the simple approximation $e^{-m_x/2}$ could be used to estimate $\mathbb{P}(\{\langle X \rangle > x\})$. In Section 4, we will show that the improvement from (3) to (4) can be critical in the accurate characterization of the tail probability.

Before we proceed, it should be mentioned that although we study the asymptotic behavior of $\mathbb{P}(\{\langle X \rangle > x\})$, as $x \to \infty$ (i.e., the x-asymptotics of the tail probability), there have been other asymptotic results derived in different limiting regimes. In particular, when $X_t^{(1)}, \ldots, X_t^{(M)}$ are *i.i.d.* processes with stationary increments, the asymptotic behavior of $\mathbb{P}(\{\langle \sum_{n=1}^{M} X_t^{(n)} \rangle > Mx\})$ as $M \to \infty$, which we call M-asymptotics, has been extensively studied. For example, in [4], Botvich and Duffield have obtained an M-asymptotic large deviation result (i.e., log-similarity), and in [13], Montgomery and De Veciana have derived a stronger M-asymptotic result using the Bahadur-Rao asymptotics. An interesting remark is that for Gaussian processes $X_t^{(n)}$, the approximation naturally following the M-asymptotic result in [4] coincides with $e^{-m_x/2}$, and the approximation suggested in [13] is equivalent to the lower bound $\mathbb{P}(\{X_{t_x} > x\}) = \Psi\left(\sqrt{m_x}\right)$. Therefore, these M-asymptotic results also support the accuracy of the approximation $e^{-m_x/2}$, and indicate that m_x is important for accurately capturing the tail probability curve. *Remark:* As will become evident shortly (e.g., by (11) in the following section), $e^{-m_x/2}$ is not very different from the lower bound $\Psi(\sqrt{m_x})$. For instance, it can easily be shown that $\log e^{-m_x/2} - \log \Psi\left(\sqrt{m_x}\right) \in O(\log x)$. This also means that both the approximations, $e^{-m_x/2}$ and $\Psi\left(\sqrt{m_x}\right)$, are supported by (4). However, since M-asymptotic results consider a limit in a different direction from that in x-asymptotics, such results cannot be extended to x-asymptotics (and *vice versa*) unless very strong properties such as uniformity of convergence can be shown.

The rest of this paper is organized as follows. In Section 2, we provide basic definitions and conditions that will be extensively referred to throughout this paper, and we derive a few preliminary results. In Section 3, we derive the main results of this paper. In Section 4, we provide a useful result (our second theorem) that characterizes the asymptotic behavior of $m_x/2$ in detail, and provides further insights into the asymptotic behavior of the tail probability. *Note:* In order to be succinct, in some of our proofs we provide limited algebraic details. Readers interested in more details are referred to our technical report [7].

2 PRELIMINARIES

Throughout the paper we assume that $\{X_t : t \geq 0\}$ is a Gaussian process with stationary increments, such that $X_0 = 0$, $\kappa = -\mathbb{E}\{X_t\}/t > 0$, and $v(t) = \mathrm{Var}\{X_t\}$ is continuous and twice differentiable. We define $\psi(t) := \log v(t)$ and $\beta := \lim_{t\to\infty} \psi(t)/\log t$ (we assume that the limit exists). It should be noted from the stationary increments property that β cannot be greater than 2. We assume that $\beta \in (0, 2)$ which covers the majority of non-trivial Gaussian processes with stationary increments.

The following conditions on $v(t)$ (and $\psi(t)$) will frequently be referred to throughout this paper.

$$\lim_{t\to\infty} t\psi'(t) = \beta. \tag{c1}$$

$$v(t) \sim St^{\beta} \quad \text{for some } S > 0 \text{ as } t \to \infty. \tag{c2}$$

$$\limsup_{t\downarrow 0} \frac{v(t)}{t^{\alpha}} < \infty \quad \text{for some } \alpha \in (0, \beta). \tag{c3}$$

$$\lim_{t\to\infty} t^2\psi''(t) = -\beta. \tag{c4}$$

Conditions (c1) and (c4) are a direct result of the definition of β, (i.e., $\beta := \lim_{t\to\infty}\psi(t)/\log t$) as long as *L'Hospital's* rule can be applied. Condition (c2) is closely related to the self-similarity of X_t, i.e., if X_t is (asymptotically) self-similar, then (c2) holds for some $\beta > 1$ (for more about self-similarity and its origination, see [11] and references therein). Also, a fairly general class of long-range dependent X_t satisfies (c2) for some $\beta > 1$. Here it should be noted that strict self-similar Gaussian processes, i.e., Fractal Brownian motion processes, constitute but a very small subset of the long-range dependent processes covered by (c2). Condition (c3) is about the behavior of $v(t)$ around $t = 0$, and will be satisfied if $v(t)$ decreases as fast as, or faster than t^α for some positive α as $t \downarrow 0$. In particular, if X_t is expressed as the integral of a stationary Gaussian process, i.e., if

$$X_t = \int_0^t \xi_\tau d\tau \tag{5}$$

where ξ_t is a stationary Gaussian process with negative mean, then (c3) always holds with any $\alpha \geq 1$, and Conditions (c1), (c2), and (c4) can be replaced by the following sufficient conditions given in terms of $C_\xi(\tau) := \mathrm{Cov}\{\xi_t, \xi_{t+\tau}\}$, the autocovariance function of ξ_t:
$\int_0^t C_\xi(\tau)\ d\tau \sim \beta\int_0^t (1-\tau/t)\ C_\xi(\tau)\ d\tau$; $2\int_0^t (t-\tau)C_\xi(\tau)d\tau \sim St^\beta$ for some $S > 0$; $tC_\xi(t) \sim (\beta^2-\beta)\int_0^t (1-\tau/t)\ C_\xi(\tau)\ d\tau$.

The parameter β in our definition is directly related to the well-known Hurst (or self-similarity) parameter H by $\beta = 2H$. Also, the empirical estimate of β has been popularly used to demonstrate self-similarity in various types of network traffic and to calculate the corresponding Hurst parameters [3, 11].

We begin our analysis by introducing the following proposition that tells us that t_x is asymptotically a linear function of x. Remember that t_x is the index at which $v(t)/(x+\kappa t)^2$ attains its maximum.

Proposition 1 *Under hypothesis (c1),*

$$t_x \sim \frac{\beta x}{(2-\beta)\kappa}.$$

`Proof of Proposition 1`: From the definition of t_x, it is clear that t_x is also the index at which $x^2 v(t)/(x+\kappa t)^2$ attains its maximum. Since $x^2 v(t)/(x+\kappa t)^2 \uparrow v(t)$ as $x \to \infty$, and since $v(t)$ increases to infinity with t, it follows that t_x should go to infinity as x increases to infinity.

Since $\log[v(t)/(x+\kappa t)^2]$ is differentiable wherever $v(t) > 0$, and since t_x should be a local maximum point of $\log[v(t)/(x+\kappa t)^2]$ that lies in the open set $\{t : v(t) > 0\}$, t_x must satisfy

$$0 = \left[\frac{d}{dt}\log\frac{v(t)}{(x+\kappa t)^2}\right]_{t=t_x} = \psi'(t_x) - \frac{2\kappa}{x+\kappa t_x}. \tag{6}$$

This equation can be rewritten as $x/(\kappa t_x) = 2/[t_x\psi'(t_x)] - 1$. Since we know that $t_x \to \infty$ as $x \to \infty$, it follows from (c1) that $2/[t_x\psi'(t_x)] - 1 \to (2-\beta)/\beta$ as $x \to \infty$. Thus, the proposition follows. `Q.E.D.`

It should be noted from the proof of Proposition 1 that even if there are multiple indices at which $v(t)/(x+\kappa t)^2$ attains its maximum, Proposition 1 holds for any choice of t_x among these indices. In fact, all the following results in this paper are independent of the choice of t_x. Also, as will be shown in the following section, under certain conditions, t_x becomes unique as x increases.

The next proposition gives us the asymptotic behavior of $m_x = (x + \kappa t_x)^2/v(t_x)$, and is derived using Proposition 1.

Proposition 2 *Under hypotheses (c1) and (c2),*

$$m_x \sim \frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}.$$

`Proof of Proposition 2`: Under (c1), it follows from Proposition 1, that $(x+\kappa t_x)^2 \sim 4x^2/(2-\beta)^2$. Also, from (c2) and Proposition 1, we know that $v(t_x) \sim S\beta^\beta x^\beta/[(2-\beta)^\beta\kappa^\beta]$. Therefore,

$$m_x = \frac{(x+\kappa t_x)^2}{v(t_x)} \sim \frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}.$$

`Q.E.D.`

For convenience, we define a stochastic process $\{Y_t^{(x)} : t \geq 0\}$ for each $x > 0$, as

$$Y_t^{(x)} := \frac{\sqrt{m_x}(X_{xt/\kappa} + xt)}{x(t+1)}.$$

From the definition of $Y_t^{(x)}$, it directly follows that for all $x > 0$ and $t \geq 0$,

$$X_t > x \quad \text{if and only if} \quad Y_{\kappa t/x}^{(x)} > \sqrt{m_x}. \tag{7}$$

Therefore, $\mathbb{P}(\{\langle X \rangle > x\})$ is equal to $\mathbb{P}(\{\langle Y^{(x)} \rangle > \sqrt{m_x}\})$. This is important because we will study the supremum distribution of X_t through $Y_t^{(x)}$. One can easily verify that $Y_t^{(x)}$ is a centered (zero mean) Gaussian process, and its variance and autocovariance can be obtained in terms of $v\,(t)$ as

$$\begin{aligned}
\sigma_{x,t}^2 \;&:=\; \mathrm{Var}\{Y_t^{(x)}\} = \frac{m_x v\,(xt/\kappa)}{x^2(t+1)^2}, \quad \text{and} \\
\mathrm{Cov}\{Y_t^{(x)}, Y_s^{(x)}\} \;&=\; \frac{m_x \mathrm{Cov}\{X_{\frac{xt}{\kappa}}, X_{\frac{xs}{\kappa}}\}}{x^2(t+1)(s+1)} = \frac{m_x\left(v\left(\frac{xt}{\kappa}\right) + v\left(\frac{xs}{\kappa}\right) - v\left(\frac{x|t-s|}{\kappa}\right)\right)}{2x^2(t+1)(s+1)},
\end{aligned}$$

where the last equality comes from $\mathrm{Cov}\{X_t, X_s\} = \left(v\,(t) + v\,(s) - v\,(|t-s|)\right)/2$, using the stationary increments property.

From the definition of m_x, note that $\sigma_{x,t}^2$ attains its maximum value of 1 at $\tilde{t}_x := \kappa t_x / x$. We now state a result from the *Extreme Value Theory* for Gaussian processes which we will use to derive our main results in the following section.

Theorem A (Theorem D.4 in [17]) *Let $\{\zeta_t : t \in [L, U]\}$ be a centered Gaussian process, and suppose that there exist constants a and γ such that $\mathbb{E}\{(\zeta_s - \zeta_t)^2\} \leq a|t-s|^\gamma$ for all $t, s \in [L, U]$. Then, there exists a constant K determined only by a and γ, such that for any $A \subset [L, U]$ and y,*

$$\mathbb{P}\left(\{\langle \zeta \rangle_A > y\}\right) \leq K(U-L)y^{2/\gamma}\Psi\left(\frac{y}{\langle \sigma \rangle_A}\right), \tag{8}$$

where $\langle \sigma \rangle_A = \sup_{t \in A} \sqrt{\mathrm{Var}\{\zeta_t\}}$.

3 MAIN RESULTS

Proposition 3 is the first main result in this section. It is an improved version of Theorem 4 in [6]. Informally, this result shows us how the tail probability $\mathbb{P}(\{\langle Y^{(x)} \rangle > \sqrt{m_x}\})$ becomes concentrated around $\tilde{t}_x$, as x increases.

Proposition 3 *Let $\delta > 0$ and $E_x := \left[\tilde{t}_x - x^{(\beta-2)/2+\delta}, \tilde{t}_x + x^{(\beta-2)/2+\delta}\right]$. Then, under hypotheses (c1)–(c4),*

$$\lim_{x\to\infty} \frac{\mathbb{P}(\{\langle Y^{(x)}\rangle_{E_x} > \sqrt{m_x}\})}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} = 1.$$

To prove the theorem, it suffices to show that for arbitrarily small $\delta > 0$

$$\lim_{x\to\infty} \frac{\mathbb{P}(\{\langle Y^{(x)}\rangle_{E_x^c} > \sqrt{m_x}\})}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} = 0, \tag{9}$$

where E_x^c is the complementary set of E_x. Hence, we only need to show the result for $0 < \delta < (2-\beta)/2$, for which the width of E_x decreases to 0 as $x \to \infty$. To prove (9), we need the following lemmas.

Lemma 1 *The following are true for all sufficiently large x.*
(i) Under hypotheses (c1)–(c3), for any $t_o > 0$,

$$\sigma_{x,t}^2 \leq \frac{8Gt_o^\beta}{S\beta^\beta(2-\beta)^{2-\beta}} \quad \textit{for } t \in [0, t_o],$$

where G is a positive number such that $v(t) \leq G(t^\beta + t^\alpha)$ for all $t \geq 0$ (from (c2) and (c3) such a $G > 0$ exists).
(ii) Under hypotheses (c1) and (c2), for any $t_o > 0$,

$$\sigma_{x,t}^2 \leq \frac{8G}{S\beta^\beta(2-\beta)^{2-\beta}t^{2-\beta}} \quad \textit{for } t \in [t_o, \infty),$$

where G here is a positive number such that $v(t) \leq Gt^\beta$ for all sufficiently large t (from (c2) such a $G > 0$ exists).
(iii) Under hypothesis (c1), for any $t_o > \beta/(2-\beta)$,

$$\log \sigma_{x,t}^2 \ \textit{is non-increasing on } [t_o, \infty).$$

(iv) Under hypothesis (c4), for any $t_o \in (0, (\beta + \sqrt{\beta/2})/(2-\beta)]$,

$$\log \sigma_{x,t}^2 \ \textit{is concave with } \frac{d^2}{dt^2}\log \sigma_{x,t}^2 \leq -\Delta_\beta \ \textit{on } \left[t_o, \frac{\beta + \sqrt{\beta/2}}{2-\beta}\right],$$

where Δ_β is a positive number determined only by β.

`Proof of Lemma 1`: (i) Let G be a positive number such that $v(t) \le G(t^\beta + t^\alpha)$ for all $t \ge 0$. From Proposition 2 we have

$$\begin{aligned}\langle \sigma_x^2 \rangle_{[0,t_o]} &\le \sup_{t\in[0,t_o]} \frac{m_x G\left(x^\beta t^\beta/\kappa^\beta + x^\alpha t^\alpha/\kappa^\alpha\right)}{x^2(t+1)^2} \\ &\le \frac{m_x G\left(t_o^\beta/\kappa^\beta + x^{\alpha-\beta} t_o^\alpha/\kappa^\alpha\right)}{x^{2-\beta}} \to \frac{4Gt_o^\beta}{S\beta^\beta(2-\beta)^{2-\beta}} \quad (\text{as } x \to \infty).\end{aligned}$$

Hence, for all sufficiently large x, $\langle \sigma_x^2 \rangle_{[0,t_o]} \le 8Gt_o^\beta/[S\beta^\beta(2-\beta)^{2-\beta}]$, and Lemma 1(i) follows.

(ii) Let G be a positive number such that $v(t) \le Gt^\beta$ for all sufficiently large t. It then follows that for all sufficiently large x,

$$\sigma_{x,t}^2 = \frac{m_x v(xt/\kappa)}{x^2(t+1)^2} \le \frac{Gm_x}{x^{2-\beta}\kappa^\beta t^{2-\beta}} \quad \text{for } t \ge t_o.$$

Note that $Gm_x/(x^{2-\beta}\kappa^\beta)$ converges to $4G/[S\beta^\beta(2-\beta)^{2-\beta}]$, as x increases. Hence, for all sufficiently large x, Lemma 1(ii) holds.

(iii) It follows from (c1) that, when x is sufficiently large, for all $t \ge t_o$ we have $x\psi'(xt/\kappa)/\kappa \le b\beta/t$, where $b = 2t_o/[\beta(t_o+1)] > 1$. From these inequalities, it can be verified that for $t \ge t_o$,

$$\frac{d}{dt}\log\sigma_{x,t}^2 = \frac{x}{\kappa}\psi'\left(\frac{x}{\kappa}t\right) - \frac{2}{t+1} \le \frac{b\beta}{t} - \frac{2}{t+1} \le 0,$$

when x is sufficiently large. Hence, Lemma 1(iii) follows.

(iv) It follows from (c4) that when x is sufficiently large, for all $t \ge t_o$,

$$\frac{x^2}{\kappa^2}\psi''\left(\frac{x}{\kappa}t\right) \le -\frac{b\beta}{t^2},$$

where $b = 2(1+\sqrt{\beta})^2/(2+\sqrt{\beta})^2 < 1$. From this inequality, one can verify that for $t \in [t_o, (\beta + \sqrt{\beta/2})/(2-\beta)]$,

$$\begin{aligned}\frac{d^2}{dt^2}\log\sigma_{x,t}^2 &= \frac{x^2}{\kappa^2}\psi''\left(\frac{x}{\kappa}t\right) + \frac{2}{(t+1)^2} \le -\frac{b\beta}{t^2} + \frac{2}{(t+1)^2} \\ &\le -\Delta_\beta := \frac{8(2-\beta)^2}{(4+\sqrt{2\beta})^2} - \frac{8\beta(\sqrt{\beta}+1)^2(2-\beta)^2}{(\sqrt{\beta}+2)^2(2\beta+\sqrt{2\beta})^2} < 0,\end{aligned}$$

when x is sufficiently large. Hence, Lemma 1(iv) follows. `Q.E.D.`

Lemma 2 *Under hypotheses (c1)–(c3), there exists a positive number* c *such that for sufficiently large* x*,*

$$\mathbb{E}\left\{\left(Y_t^{(x)} - Y_s^{(x)}\right)^2\right\} \leq c|t-s|^\alpha.$$

`Proof of lemma 2`: Let G be a positive number such that $v(t) \leq G(t^\beta + t^\alpha)$ for all $t \geq 0$. If we assume that $t > s$, then after some algebraic manipulation we get the following inequality:

$$\begin{aligned}
\mathbb{E}^{\frac{1}{2}}\left\{\left(Y_t^{(x)} - Y_s^{(x)}\right)^2\right\} &= \mathbb{E}^{\frac{1}{2}}\left\{\left(\frac{\sqrt{m_x}(X_{xt/\kappa}+xt)}{x(t+1)} - \frac{\sqrt{m_x}(X_{xs/\kappa}+xs)}{x(s+1)}\right)^2\right\} \\
&\leq \frac{\sqrt{m_x G}(t-s)^{\alpha/2}}{x^{1-\beta/2}\kappa^{\beta/2}}\left(\frac{t^{(\beta-\alpha)/2} + x^{(\alpha-\beta)/2}/\kappa^{(\alpha-\beta)/2}}{t+1}\right. \\
&\quad \left. + \frac{t^{1-\alpha/2}\left(s^{\beta/2} + s^{\alpha/2}x^{(\alpha-\beta)/2}/\kappa^{(\alpha-\beta)/2}\right)}{(t+1)(s+1)}\right). \qquad (10)
\end{aligned}$$

Note that $\sqrt{m_x G}/(x^{1-\beta/2}\kappa^{\beta/2})$ converges to $\sqrt{4G/S\beta^\beta(2-\beta)^{2-\beta}}$ as $x \to \infty$ (from Proposition 2). Further note that both $(t^{(\beta-\alpha)/2} + (x/\kappa)^{(\alpha-\beta)/2})/(t+1)$ and $t^{1-\alpha/2}(s^{\beta/2} + s^{\alpha/2}(x/\kappa)^{(\alpha-\beta)/2})/[(t+1)(s+1)]$ can be bounded by a positive number for all large x. Thus, there exists a $c > 0$ such that $\sqrt{c}|t-s|^{\alpha/2}$ bounds the last expression in (10) for all large x, and the lemma follows. `Q.E.D.`

We are now ready to prove Proposition 3.

`Proof of Proposition 3`: Define

$$T := \min\left\{\frac{\beta - \sqrt{\beta/2}}{2-\beta}, \left(\frac{S\beta^\beta(2-\beta)^{2-\beta}}{16G}\right)^{1/\beta}\right\}$$

and

$$J := \left\lceil \max\left\{\frac{\beta + \sqrt{\beta/2}}{2-\beta}, \left(\frac{16G}{S\beta^\beta(2-\beta)^{2-\beta}}\right)^{1/(2-\beta)}\right\}\right\rceil + 1,$$

where $\lceil w \rceil$ denotes the smallest integer larger than or equal to w. Now, we will apply Theorem A to $Y_t^{(x)}$ over closed intervals $[0, T]$, $[T, J]$, $[J, J+1]$, $[J+1, J+2]$, $\ldots$ for all sufficiently large x, and derive an asymptotic upper bound on $\mathbb{P}(\{\langle Y^{(x)}\rangle_{E_x^c} > \sqrt{m_x}\})$. This bound together with the lower bound

$\Psi(\sqrt{m_x})$ will result in (9). Henceforth, consider only sufficiently large x for which Lemma 1 (with explicitly or implicitly given values of t_o) and Lemma 2 hold. Also, let K denote the coefficient in (8) determined only by c and α in Lemma 2, and G be a positive number such that $v\,(t) \leq G(t^\beta + t^\alpha)$ for all $t \geq 0$ and $v\,(t) \leq Gt^\beta$ for all sufficiently large t.

From Theorem A and Lemma 2 ((i) with $t_o = T$), we have

$$\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[0,T]} > \sqrt{m_x}\right\}\right) \leq KTm_x^{1/\alpha}\Psi\left(\sqrt{2m_x}\right).$$

This result, together with Proposition 2 and a well-known bound for $\Psi(z)$ [9], i.e.,

$$\frac{1-z^{-2}}{\sqrt{2\pi}}z^{-1}e^{-z^2/2} \leq \Psi\,(z) \leq \frac{1}{\sqrt{2\pi}}z^{-1}e^{-z^2/2} \quad \text{for all } z > 0, \tag{11}$$

implies that

$$\begin{aligned} \log \mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[0,T]} > \sqrt{m_x}\right\}\right) &\leq \log \frac{KTm_x^{(2-\alpha)/2\alpha}e^{-m_x}}{2\sqrt{\pi}} \\ &\sim -\frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} \quad (\text{as } x \to \infty). \end{aligned} \tag{12}$$

On the other hand, remember that

$$\log \mathbb{P}\left(\left\{\langle Y^{(x)}\rangle > \sqrt{m_x}\right\}\right) \geq \log \Psi(\sqrt{m_x}) \sim -\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}},$$

where the similarity (as $x \to \infty$) is again from Proposition 2 and (11). By comparing (12) and (13), we get

$$\frac{\mathbb{P}\left(\{\langle Y^{(x)}\rangle_{[0,T]} > \sqrt{m_x}\}\right)}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} \to 0 \quad \text{as } x \to \infty. \tag{13}$$

We next show a similar result for the interval $[J, \infty)$. We start with the relation

$$\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J,\infty)} > \sqrt{m_x}\right\}\right) \leq \sum_{n=0}^{\infty} \mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J+n,J+n+1]} > \sqrt{m_x}\right\}\right).$$

If we define $D := 8G/S\beta^\beta(2-\beta)^{2-\beta}$, then from Theorem A, Lemma 1 ((ii) with $t_o = J$), and Lemma 2, it follows that

$$\log \mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J+n,J+n+1]} > \sqrt{m_x}\right\}\right) \leq Km_x^{1/\alpha}\Psi\left(\sqrt{\frac{m_x(J+n)^{2-\beta}}{D}}\right)$$

$$\leq K\sqrt{D}m_x^{(2-\alpha)/2\alpha}\frac{e^{-m_x(J+n)^{2-\beta}/2D}}{\sqrt{2\pi(J+n)^{2-\beta}}},$$

where the last inequality is from (11). Consequently,

$$\begin{aligned}\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J,\infty)} > \sqrt{m_x}\right\}\right) &\leq K\sqrt{D}m_x^{(2-\alpha)/2\alpha}\sum_{n=0}^{\infty}\frac{e^{-m_x(J+n)^{2-\beta}/2D}}{\sqrt{2\pi(J+n)^{2-\beta}}} \\ &\leq K\sqrt{D}m_x^{(2-\alpha)/2\alpha}\int_{J-1}^{\infty}e^{-m_x y^{2-\beta}/2D}dy. \quad (14)\end{aligned}$$

Further, by replacing y in the integral with $z = m_x y^{2-\beta}/2D$, it can easily be shown that there exists a polynomial $p(z)$ such that

$$\int_{J-1}^{\infty}e^{-m_x y^{2-\beta}/2D}dy \leq p(m_x)e^{-m_x(J-1)^{2-\beta}/2D}, \qquad \text{for all large } x.$$

Therefore, from (14) and the definitions of J and D, one can see that $\log\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J,\infty)} > \sqrt{m_x}\right\}\right)$ will decrease at least as fast as $-4\kappa^\beta x^{2-\beta}/[S\beta^\beta(2-\beta)^{2-\beta}]$, as $x\to\infty$. This result, together with (13), implies that

$$\frac{\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[J,\infty)} > \sqrt{m_x}\right\}\right)}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} \to 0 \quad \text{as } x\to\infty. \quad (15)$$

Now, given (13) and (15), to complete the proof (i.e., to show (9)), we need only to show that

$$\lim_{x\to\infty}\frac{\mathbb{P}(\{\langle Y^{(x)}\rangle_{[T,J]\cap E_x^c} > \sqrt{m_x}\})}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} = 0. \quad (16)$$

From Proposition 1, note that $\tilde{t}_x \to \beta/(2-\beta)$, as $x\to\infty$. Therefore, we have $E_x \subset [T, (\beta+\sqrt{\beta/2})/(2-\beta)]$ for all sufficiently large x (remember that we assume $\delta < (2-\beta)/2$, and hence, the size of interval E_x decreases to 0). From Lemma 1 (iii) and (iv), $\log\sigma_{x,t}^2$ is non-increasing on $[(\beta+\sqrt{\beta/2})/(2-\beta), J]$ and concave with $d^2\log\sigma_{x,t}^2/dt^2 \leq -\Delta_\beta$ on $[T, (\beta+\sqrt{\beta/2})/(2-\beta)]$. All these facts together with the fact that $[\log\sigma_{x,t}^2]_{t=\tilde{t}_x} = \log\langle\sigma_x^2\rangle = 0$ (from the definition of m_x and $Y_t^{(x)}$) and $[\frac{d}{dt}\log\sigma_{x,t}^2]_{t=\tilde{t}_x} = 0$ (since $\sigma_{x,t}^2$ attains its maximum at $t=\tilde{t}_x$), imply that

$$\langle\sigma_x^2\rangle_{[T,J]\cap E_x^c} \leq \exp\left[-\Delta_\beta x^{\beta-2+2\delta}/2\right] \leq \frac{2}{2+\Delta_\beta x^{\beta-2+2\delta}} \quad \text{for all large } x.$$

By applying Theorem A to $\{Y_t^{(x)} : t \in [T, J]\}$, we have

$$\begin{aligned}
\mathbb{P}\left(\left\{\langle Y^{(x)}\rangle_{[T,J]\cap E_x^c} > \sqrt{m_x}\right\}\right) &\leq K(J-T)m_x^{1/\alpha}\ \Psi\left(\sqrt{\frac{m_x(2+\Delta_\beta x^{\beta-2+2\delta})}{2}}\right) \\
&\leq \frac{K(J-T)m_x^{(2-\alpha)/2\alpha}}{\sqrt{\pi(2+\Delta_\beta x^{\beta-2+2\delta})}}\ e^{-m_x(2+\Delta_\beta x^{\beta-2+2\delta})/4},
\end{aligned}$$

where the last inequality is from (11). Therefore, it follows that

$$\begin{aligned}
\log \frac{\mathbb{P}\left(\{\langle Y^{(x)}\rangle_{[T,J]} > \sqrt{m_x}\}\right)}{\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})} &\leq \log \frac{\mathbb{P}\left(\{\langle Y^{(x)}\rangle_{[T,J]} > \sqrt{m_x}\}\right)}{\Psi\left(\sqrt{m_x}\right)} \\
&\leq \log\left(\frac{\sqrt{2}K(J-T)m_x^{1/\alpha}\ e^{-m_x(\Delta_\beta x^{\beta-2+2\delta})/4}}{\sqrt{(2+\Delta_\beta x^{\beta-2+2\delta})}(1-m_x^{-1})}\right) \\
&\qquad\qquad \text{(from (11))} \\
&\sim -\frac{\Delta_\beta \kappa^\beta x^{2\delta}}{S\beta^\beta(2-\beta)^{2-\beta}} \quad \text{as } x \to \infty \\
&\qquad\qquad \text{(from Proposition 2)},
\end{aligned}$$

and (16) follows. Q.E.D.

If we rewrite Proposition 3 as

$$\lim_{x\to\infty} \mathbb{P}(\{\langle Y^{(x)}\rangle_{E_x} > \sqrt{m_x} | \langle Y^{(x)}\rangle > \sqrt{m_x}\}) = 1, \tag{17}$$

then its implication becomes more evident. Note that if we choose δ small enough, the size of the interval E_x around $\tilde{t}_x$ decreases to 0 as x increases. Hence, (17) tells us that the rarer the event becomes, the more the event concentrates around the most likely time $\tilde{t}_x$. Therefore, Proposition 3 explains why the lower bound $\mathbb{P}(\{X_{t_x} > x\}) = \mathbb{P}(\{Y_{\tilde{t}_x}^{(x)} > \sqrt{m_x}\})$ accurately bounds the tail probability $\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\}) = \mathbb{P}(\{\langle X\rangle > x\})$. In our previous research [5], we have numerically investigated the accuracy of this lower bound as an approximation to $\mathbb{P}(\{\langle X\rangle > x\})$ under more restrictive conditions; i.e., when X_t can be expressed as (5) where ξ_t is a stationary Gaussian process with negative mean and *absolutely integrable autocovariance function* $C_\xi(\tau)$. There, we have found that the lower bound is fairly accurate and matches the curve of $\mathbb{P}(\{\langle X\rangle > x\})$ over a wide range of values of x. Further, in [5, 6], it has been observed that $e^{-m_x/2}$ provides an asymptotic upper bound to

$\mathbb{P}(\{\langle X\rangle > x\})$ and approximates the tail probability just as accurately as the lower bound. Note that it can easily be verified that the absolute intergability of $C_\xi(\tau)$ implies that the corresponding value of β should be equal to 1. In other words, our previous results in [5, 6] do not hold for long-range dependent X_t. However, since Proposition 3 is valid for a more general class of processes X_t, including those exhibiting long-range dependence, we expect that both the lower bound $\Psi(\sqrt{m_x})$ and the asymptotic upper bound $e^{-m_x/2}$ will accurately approximate the tail probability, even for long-range dependent X_t. In the next theorem, we provide an important asymptotic property of $\mathbb{P}(\{\langle X\rangle > x\})$ that supports this conjecture.

Theorem 1 *Under hypotheses (c1)–(c3),*

$$\begin{aligned} -\infty &< \liminf_{x\to\infty} \frac{1}{\log x}\left(\log \mathbb{P}(\{\langle X\rangle > x\}) + \frac{m_x}{2}\right) \\ &\le \limsup_{x\to\infty} \frac{1}{\log x}\left(\log \mathbb{P}(\{\langle X\rangle > x\}) + \frac{m_x}{2}\right) < \infty. \end{aligned}$$

`Proof of Theorem 1`: From (7), (13), and (15), it suffices to show the theorem with $\mathbb{P}(\{\langle X\rangle > x\})$ replaced by either $\mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})$ or $\mathbb{P}(\{\langle Y^{(x)}\rangle_{[T,J]} > \sqrt{m_x}\})$, where T and J are defined in the proof of Proposition 3.

The lim-inf part of the theorem directly follows from the lower bound $\Psi(\sqrt{m_x})$. In other words, from (11), (13), and Proposition 2, we have

$$\log \mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\}) + \frac{m_x}{2} \ge \log \Psi\left(\sqrt{m_x}\right) + \frac{m_x}{2} \sim -\frac{2-\beta}{2}\log x.$$

Therefore, it follows that

$$\liminf_{x\to\infty} \frac{1}{\log x}\left(\log \mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\}) + \frac{m_x}{2}\right) > -\infty.$$

To prove the lim-sup part of the theorem, we apply Theorem A to $\{Y_t^{(x)} : t \in [T,J]\}$. From the definition of $Y_t^{(x)}$ and Proposition 1, we already know that $\langle \sigma_x^2\rangle_{[T,J]} = 1$ for all large x. Again, let K denote the coefficient in (8) determined only by c and α in Lemma 2. Then, Theorem A and Lemma 2 imply that

$$\begin{aligned} \log \mathbb{P}(\{\langle Y^{(x)}\rangle_{[T,J]} > \sqrt{m_x}\}) + \frac{m_x}{2} &\le \log K(J-T)m_x^{1/\alpha}\Psi\left(\sqrt{m_x}\right) + \frac{m_x}{2} \\ &\sim \frac{(2-\alpha)(2-\beta)}{2\alpha}\log x. \end{aligned}$$

Therefore, it follows that

$$\limsup_{x\to\infty} \frac{1}{\log x}\left(\log \mathbb{P}(\{\langle Y^{(x)}\rangle_{[T,J]} > \sqrt{m_x}\}) + \frac{m_x}{2}\right) < \infty.$$

Q.E.D.

Note that (4) in the introduction is a compact form representation of Theorem 1. The theorem suggests that $-m_x/2$ is a good estimate of $\log \mathbb{P}(\{\langle X\rangle > x)\} = \log \mathbb{P}(\{\langle Y^{(x)}\rangle > \sqrt{m_x}\})$ in the sense that the error could *at most* increase on the order of $\log x$. In other words,

$$\mathbb{P}(\{\langle X\rangle > x\}) = \exp\left[-\frac{m_x}{2} + r(x)\right], \tag{18}$$

where $r(x) := \log \mathbb{P}(\{\langle X\rangle > x\}) + m_x/2 \in O(\log x)$.

We next relate our result with existing Large Deviation results. Note that the leading term of m_x by itself satisfies the Large Deviation relation (3), i.e., from Proposition 2 and Theorem 1, we have

$$\log \mathbb{P}(\{\langle X\rangle > x\}) \sim -\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}. \tag{19}$$

If we define $R(x) := \log \mathbb{P}(\{\langle X\rangle > x\}) + 2\kappa^\beta x^{2-\beta}/S\beta^\beta(2-\beta)^{2-\beta}$, then the tail probability can also be written as

$$\mathbb{P}(\{\langle X\rangle > x\}) = \exp\left[-\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} + R(x)\right]. \tag{20}$$

Further, it follows from (19) that $R(x) \in o(x^{2-\beta})$, where $o(f(x))$ denotes the set of functions $g(x)$ such that for any $\epsilon > 0$, $|g(x)| \le \epsilon|f(x)|$ eventually. Since $o(x^{2-\beta})$ is a much larger set than $O(\log x)$, as mentioned in Section 1, (4) characterizes the asymptotic tail behavior in much more detail than (19) and therefore significantly improves upon the resolution of (19). This tells us that an approximation based on (18), i.e.,

$$\mathbb{P}(\{\langle X\rangle > x\}) \approx \exp\left[-\frac{m_x}{2}\right] \tag{21}$$

should be significantly more accurate than

$$\mathbb{P}(\{\langle X\rangle > x\}) \approx \exp\left[-\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}\right], \tag{22}$$

the approximation that is naturally suggested by (20).

From a practical viewpoint, it may be important to know whether t_x is unique or not. For example, when $e^{-m_x/2}$ has to be computed to approximate $\mathbb{P}(\{\langle X\rangle > x\})$, its accuracy and computation time will be determined by how fast and accurately we can calculate m_x. The following proposition guarantees the accurate and fast computation of t_x and m_x, for large x.

Proposition 4 *Under hypotheses (c1) and (c4), for all sufficiently large x, $\log[v(t)/(x+\kappa t)^2]$ is strictly concave on $[(\beta-\sqrt{\beta/2})x/(2-\beta)\kappa, (\beta+\sqrt{\beta/2})x/(2-\beta)\kappa]$, and there is a unique index t_x where $v(t)/(x+\kappa t)^2$ attains its maximum m_x.*

`Proof of Proposition 4`: From the direct relation between $\sigma^2_{x,t}$ and $v(t)/(x+\kappa t)^2$, it suffices to show that $\log\sigma^2_{x,t}$ is strictly concave on $[(\beta-\sqrt{\beta/2})/(2-\beta), (\beta+\sqrt{\beta/2})/(2-\beta)]$, and that there is a unique index $\tilde{t}_x$ where it attains its maximum.

From Lemma 1(iv), when x is sufficiently large,

$$\frac{d^2}{dt^2}\log\sigma^2_{x,t} < 0 \quad \text{for } t \in \left[(\beta-\sqrt{\beta/2})/(2-\beta), (\beta+\sqrt{\beta/2})/(2-\beta)\right],$$

which proves the concavity of $\log\sigma^2_{x,t}$. Now, from Proposition 1, $\log\sigma^2_{x,t}$ attains its maximum value in $[(\beta-\sqrt{\beta/2})/(2-\beta), (\beta+\sqrt{\beta/2})/(2-\beta)]$ for all sufficiently large x. However, since $\log\sigma^2_{x,t}$ is strictly concave on this interval when x is large enough, there cannot be more than one index in the interval over which $\log\sigma^2_{x,t}$ takes its maximum value. `Q.E.D.`

Proposition 4 tells us that *when x is large, t_x and m_x can be computed by performing a simple local search algorithm starting at $\beta x/[(2-\beta)\kappa]$*. Although this proposition is valid only for large enough x, *according to our numerical studies, local search algorithms usually find t_x and m_x accurately within a small number of iterations*, even for fairly small values of x. This is because $v(t)/(x+\kappa t)^2$ is usually of a distinctly unimodal shape even for fairly small values of x.

In the next section we study the asymptotic properties of m_x in more detail, and the effect of its secondary terms on the asymptotic behavior of $\mathbb{P}(\{\langle X\rangle > x\})$.

4 THE IMPACT OF $\mathrm{Var}\{X_t\}$ ON THE ASYMPTOTIC BEHAVIOR OF m_x AND $\mathbb{P}(\{\langle X\rangle > x\})$

In this section we will introduce our second theorem, which relates the asymptotic behavior of $\mathrm{Var}\{X_t\}$ to that of m_x and $\mathbb{P}(\{\langle X\rangle > x\})$. We begin with a simple example of Fractal Brownian motion processes, a well-known class of self-similar processes [16].

4.1 Fractal Brownian Motion Process

The standard (normalized) Fractal Brownian motion process $\{B_t^{(H)} : t \geq 0\}$ with Hurst parameter $H \in (0,1)$ is a centered Gaussian process with stationary increments that possesses the following properties [16]: (i) $B_0^{(H)} = 0$; (ii) $\mathrm{Var}\{B_t^{(H)}\} = t^{2H}$; and (iii) $B_t^{(H)}$ is sample path continuous.

We first study the supremum distribution of $X_t := SB_t^{(H)} - \kappa t$ which is often called Fractal Brownian motion with negative linear drift.

From the above properties of Fractal Brownian motion processes, one can easily verify that X_t satisfies all conditions (c1)–(c4) with $\beta = 2H$. Also, in this case we can compute m_x explicitly as

$$m_x = \frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}. \tag{23}$$

Therefore, for Fractal Brownian motion processes with negative linear drift, (19) can be strengthened (by substituting (23) into (4)) to

$$R(x) = \log \mathbb{P}(\{\langle X\rangle > x\}) + \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} \in O(\log x). \tag{24}$$

In other words, for Fractal Brownian motion with negative linear drift, the approximation based on the Large Deviation techniques can in fact be better supported by Theorem 1. Of course, note that in this case (21) and (22) result in the same approximation. Here, it should also be mentioned that the exact queue asymptotics with Fractal Brownian Motion inputs have been provided in [15].

It should be noted here that the reason why (19) can be strengthened to (24) for Fractal Brownian motion processes with negative linear drift, is that $-m_x/2$ is not different from the leading term captured by (19). In fact, since we know that $r(x) \in O(\log x)$, one can easily see that (24) holds if and only if

$$r(x) - R(x) = \frac{m_x}{2} - \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} \in O(\log x). \tag{25}$$

Therefore, if (25) holds, the approximation (22) can be supported by (24) just as strongly as the approximation (21) is supported by (4).

We next provide our third theorem which shows, with more precision than Proposition 2, how the asymptotic behavior of $\mathrm{Var}\{X_t\}$ impacts the asymptotic behavior of m_x. This theorem results in a corollary, which provides a sufficient condition for (25) to hold. To state the Theorem succinctly, we define a function $h(t) := v(t)/(St^\beta) - 1$.

Theorem 2 *Under hypotheses (c1) and (c2), if there exists a positive constant C such that*

$$t|h'(t)| \le C|h(t)| \quad \textit{for all large } t, \tag{26}$$

then

$$\frac{m_x}{2} - \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} \sim -\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} h(t_x).$$

`Proof of Theorem 2`: We first obtain several simple asymptotic relations. Let $\check{t}_x := \beta x/(2-\beta)\kappa$, then (6) can be rewritten as

$$t_x - \check{t}_x = \frac{2t_x(t_x\psi'(t_x) - \beta)}{(2-\beta)t_x\psi'(t_x)}. \tag{27}$$

Further, it follows from the definition of $h(t)$ that

$$t\psi'(t) - \beta = \frac{St^{\beta+1}h'(t)}{v(t)}. \tag{28}$$

Therefore, from (c1), (c2), (27), (28), and Proposition 1, we have

$$t_x - \check{t}_x = \frac{2St_x^{\beta+2}h'(t_x)}{(2-\beta)v(t_x)\,t_x\psi'(t_x)} \sim \frac{2\beta x^2}{(2-\beta)^3\kappa^2} h'(t_x), \tag{29}$$

and using the Mean Value Theorem we get

$$t_x^\beta - \check{t}_x^\beta \sim \frac{2\beta^{\beta+1}x^{\beta+1}}{(2-\beta)^{\beta+2}\kappa^{\beta+1}} h'(t_x). \tag{30}$$

Now, from the definition of $\check{t}_x$ and t_x, it follows that $4\kappa^\beta x^{2-\beta}/[S\beta^\beta(2-\beta)^{2-\beta}] = (x+\kappa\check{t}_x)^2/S\check{t}_x^\beta$ and $m_x = (x+\kappa t_x)^2/v(t_x)$. Hence, we can write

$$\begin{aligned} m_x - \frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} &= \frac{(x+\kappa t_x)^2}{v(t_x)} - \frac{(x+\kappa\check{t}_x)^2}{S\check{t}_x^\beta} \\ &= \frac{(x+\kappa t_x)^2(St_x^\beta - v(t_x))}{St_x^\beta v(t_x)} + \frac{(x+\kappa t_x)^2(\check{t}_x^\beta - t_x^\beta)}{S\check{t}_x^\beta t_x^\beta} \\ &\quad + \frac{\kappa(t_x-\check{t}_x)(2x+\kappa(t_x+\check{t}_x))}{S\check{t}_x^\beta}. \end{aligned} \tag{31}$$

Moreover, we have

$$\frac{(x+\kappa t_x)^2(St_x^\beta - v(t_x))}{St_x^\beta v(t_x)} \sim -\frac{4\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} h(t_x) \tag{32}$$

(from (c2), Proposition 1, and the definition of $h(t)$),

$$\frac{(x+\kappa t_x)^2(\check{t}_x^\beta - t_x^\beta)}{S\check{t}_x^\beta t_x^\beta} \sim -\frac{8\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{3-\beta}} t_x h'(t_x) \tag{33}$$

(from Proposition 1, (30), and the definition of $\check{t}_x$), and

$$\frac{\kappa(t_x-\check{t}_x)(2x+\kappa(t_x+\check{t}_x))}{S\check{t}_x^\beta} \sim \frac{8\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{3-\beta}} t_x h'(t_x) \tag{34}$$

(from Proposition 1, (29), and the definition of $\check{t}_x$). Note that (26), (33), and (34) imply that

$$\frac{(x+\kappa t_x)^2(\check{t}_x^\beta - t_x^\beta)}{S\check{t}_x^\beta t_x^\beta} + \frac{\kappa(t_x-\check{t}_x)(2x+\kappa(t_x+\check{t}_x))}{S\check{t}_x^\beta} \in o\left(x^{2-\beta}h(t_x)\right). \tag{35}$$

Therefore, it follows from (31), (32), and (35), that

$$\frac{m_x}{2} - \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} \sim -\frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} h(t_x).$$

Q.E.D.

Corollary 1 *Under hypotheses (c1)–(c3), if there exists a positive constant C that satisfies (26) and if $h(t) \in O(t^{\beta-2}\log t)$, then (25) holds.*

`Proof of Corollary 1`: Assuming that $h(t) \in O(t^{\beta-2}\log t)$, then there exists a positive constant c such that $|h(t)| \leq ct^{\beta-2}\log t$ for all large t. Hence, it follows from Theorem 2 that, for all large x,

$$\begin{aligned} |r(x) - R(x)| &= \left|\frac{m_x}{2} - \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}}\right| \\ &\leq \frac{3c\kappa^\beta x^{2-\beta} t_x^{\beta-2}\log t_x}{S\beta^\beta(2-\beta)^{2-\beta}}. \end{aligned}$$

Since $t_x^{\beta-2}\log t_x \sim \beta^{\beta-2}x^{\beta-2}\log x/[(2-\beta)^{\beta-2}\kappa^{\beta-2}]$ from Proposition 1, this implies that $|r(x) - R(x)| \leq 4c\kappa^2 \log x/S\beta^2$ for all large x. Hence, we have $r(x) - R(x) \in O(\log x)$. `Q.E.D.`

The additional condition beyond (c1)–(c3) that Theorem 2 requires, is not a restrictive condition since we can usually find a constant $C > 0$ that satisfies (26) (as will be illustrated in the following section). Therefore, Theorem 2 not only results in a sufficient condition (Corollary 1) for (25) to hold, but also tells us in considerable generality *when (25) will not hold.* Let $u_t := v(t) - St^\beta$. From the definition, u_t consists of all terms of $v(t)$ other than the leading term St^β. Further assume that $u_t \sim ct^\gamma$ for some $\gamma \in (0, \beta)$ and $c \neq 0$. From the definitions of $h(t)$ and u_t, we can then see that $h(t) = u_t/St^\beta \sim ct^{\gamma-\beta}/S$. Hence, provided that (26) holds for some $C > 0$, Theorem 2 implies that

$$\frac{m_x}{2} - \frac{2\kappa^\beta x^{2-\beta}}{S\beta^\beta(2-\beta)^{2-\beta}} = r(x) - R(x) \sim -\frac{2c\kappa^{2\beta-\gamma}x^{2+\gamma-2\beta}}{S^2\beta^{2\beta-\gamma}(2-\beta)^{2+\gamma-2\beta}}. \quad (36)$$

In other words, Theorem 2 tells us that asymptotically, the more slowly growing the additional term u_t is, the more $-m_x/2$ behaves like $-2\kappa^\beta x^{2-\beta}/[S\beta^\beta(2-\beta)^{2-\beta}]$. In this sense, the Fractal Brownian motion process with negative linear drift that we have just studied, is an extreme case, where $u_t = 0$ for all t and (25) trivially holds. Hence, we next consider situations when u_t is not identically equal to 0, and show that *(25) does not hold in many cases.*

4.2 Other Gaussian Processes with stationary increments

From (36), if $\gamma \in (2\beta - 2, \beta)$, then $r(x) - R(x)$ increases as a power of x, and since $r(x) \in O(\log x)$, it follows that

$$R(x) \sim \frac{2c\kappa^{2\beta-\gamma}x^{2+\gamma-2\beta}}{S^2\beta^{2\beta-\gamma}(2-\beta)^{2+\gamma-2\beta}}. \quad (37)$$

Therefore, if $\gamma \in (2\beta - 2, \beta)$, then (24) will not hold. Through a simple example, we now show that γ can, in fact, be greater than $2\beta - 2$.

Consider a stationary Gaussian process ξ_t with mean and autocovariance given by

$$\mathbb{E}\{\xi_t\} = -\kappa \quad \text{and} \tag{38}$$

$$\mathrm{Cov}\{\xi_t, \xi_{t+\tau}\} = \frac{S\beta(\beta-1)}{2(|\tau|+1)^{2-\beta}} \quad \text{with } 1 < \beta < \frac{3}{2}. \tag{39}$$

If we define $\{X_t : t \geq 0\}$ by (5), then one can easily verify that X_t is a Gaussian process with stationary increments that satisfies (c1)–(c3), and that

$$\mathbb{E}\{X_t\} = -\kappa t \quad \text{and}$$

$$\mathrm{Var}\{X_t\} = 2\int_0^t \frac{S\beta(\beta-1)(t-\tau)}{2(\tau+1)^{2-\beta}} d\tau = S((t+1)^\beta - 1) - S\beta t. \tag{40}$$

From (40), it can easily be verified that $u_t = v(t) - St^\beta \sim -S\beta t$ and that $th'(t) \sim (1-\beta)h(t)$ (a sufficient condition for (26) to hold for some C). Therefore, we actually have $\gamma = 1 > 2\beta - 2$ in this case, and it follows from (37), that $R(x) \sim -2\kappa^{2\beta-1}x^{3-2\beta}/[S\beta^{2\beta-2}(2-\beta)^{3-2\beta}]$. This means that the approximation (22) will result in an error (for large x) roughly on the order of $\exp[-2\kappa^{2\beta-1}x^{3-2\beta}/(S\beta^{2\beta-2}(2-\beta)^{3-2\beta})]$, and could result in significantly overestimating $\mathbb{P}(\{\langle X \rangle > x\})$.

To illustrate the error involved in (22), in Figure 2 we show numerical results of approximations $\exp[-2\kappa^\beta x^{2-\beta}/(S\beta^\beta(2-\beta)^{2-\beta})]$ and $e^{-m_x/2}$ for $\mathbb{P}(\{\langle X \rangle > x\})$, when the stationary Gaussian process ξ_t in (5) is specified by (38) and (39) with $S = 4$, $\kappa = 1$, and $\beta = 5/4$. As one can see from the figure, the two approximations diverge from each other very quickly (their difference increases by nearly 2 orders of magnitude while x increases from 10 to 90). Since we know from (4) that the tail probability $\mathbb{P}(\{\langle X \rangle > x\})$ does not rapidly diverge from $e^{-m_x/2}$, this numerical result shows that the approximation (22) based on (19) will eventually (as x increases) result in a serious error.

As illustrated through the preceding simple example, the additional term u_t can be quite arbitrary. Now consider another example: a fluid queue model of a high-speed multiplexer, with an infinite buffer and a constant link rate μ, serving M independent stationary Gaussian inputs with instantaneous input rate $\lambda_t^{(n)} (n = 1, 2, \ldots, M)$; μ represents the bandwidth of the output link, and

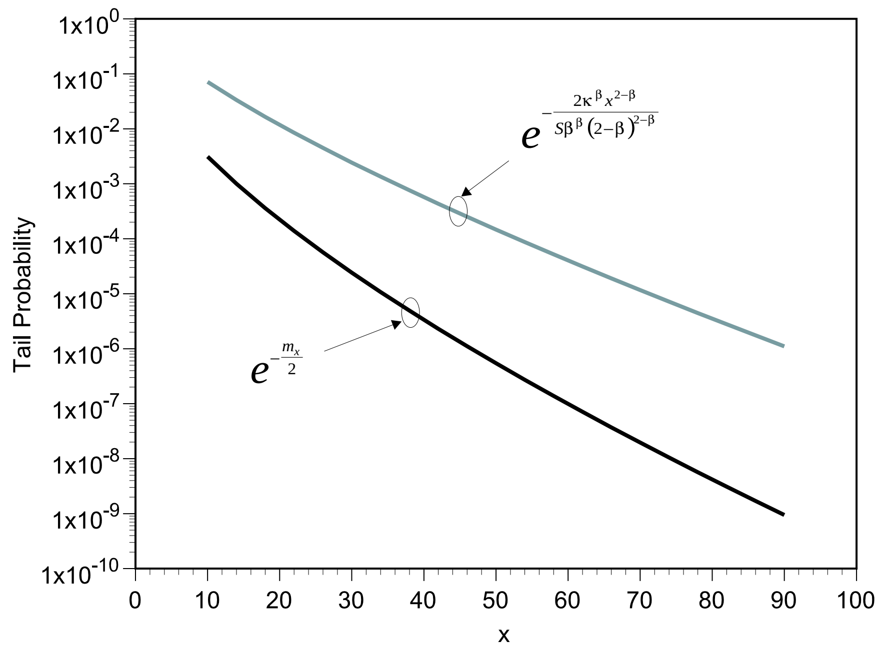

Figure 2: Two approximations based on (4) and (19) for $\mathbb{P}(\{\langle X\rangle > x\})$, when X_t is defined by (5), and ξ_t is a stationary Gaussian process with $\mathbb{E}\{\xi_t\} = -1$ and $\mathrm{Cov}\{\xi_t, \xi_{t+\tau}\} = 5(|\tau|+1)^{-3/4}/8$.

$\lambda_t^{(n)} (n = 1, 2, \ldots, M)$ models the packet arrivals from the M different streams being multiplexed. Hence, the net amount of input during the interval $(s, t]$ is $\mathrm{N}_t - \mathrm{N}_s = \int_s^t \sum_{n=1}^M \lambda_\tau^{(n)} - \mu d\tau$.

Also, from (1), the tail $\mathbb{P}(\{Q > x\})$ of the steady-state queue length distribution is the same as $\mathbb{P}(\{\langle X\rangle > x\})$, where $\{X_t : t \geq 0\}$ is defined by $X_t := \int_{-t}^0 \sum_{n=1}^M \lambda_\tau^{(n)} - \mu d\tau$. Note that the variance of X_t can be expressed in terms of the autocovariance $C_n(\tau) := \mathrm{Cov}\{\lambda_t^{(n)}, \lambda_{t+\tau}^{(n)}\}$ of M input processes as

$$v(t) = \sum_{n=1}^{M} 2 \int_0^t (t-\tau) C_n(\tau) d\tau. \tag{41}$$

In other words, $v(t)$ is composed of M terms, each of which is determined by the autocovariance of the corresponding input process. If we assume that $2\int_0^t (t-\tau) C_n(\tau) d\tau \sim S_n t^{\beta_n}$ for some $S_n > 0$ and $\beta_n \in (0, 2)$, one can easily see that $v(t) \sim \hat{S}_1 t^{\hat{\beta}_1}$, where $\hat{\beta}_1 = \max\{\beta_1, \beta_2, \ldots, \beta_M\}$ and $\hat{S}_1 = \sum_{\{n:\beta_n = \hat{\beta}_1\}} S_n$. Therefore, if the values of β_n are not identical, the leading term $\hat{S}_1 t^{\hat{\beta}_1}$ of $v(t)$ will capture only the terms in (41) that increase on the order of $t^{\hat{\beta}_1}$, and hence

u_t is likely to increase on the order of $t^{\hat{\beta}_2}$, where $\hat{\beta}_2$ is the second largest value among $\{\beta_1, \beta_2, \ldots, \beta_M\}$. If $\hat{\beta}_2$ is greater than $2\hat{\beta}_1 - 2$ in this case, then from (37), $R(x)$ will grow on the order of $x^{2+\hat{\beta}_2-2\hat{\beta}_1}$, and the approximation (22) will be poor for large x.

Empirical studies on various types of network traffic suggest that the traffic is composed of the aggregation of statistically very diverse sources that may exhibit different kinds of long-range dependence. As illustrated in the previous example, under this heterogeneity, the Large Deviation result (19) may not be precise enough to capture the queueing behavior of the network traffic, and an approximation like (22) should be used with caution.

In contrast, (4) considerably improves the resolution of (19), and naturally leads to the approximation (21), which, if it does diverge, does not diverge rapidly from $\mathbb{P}(\{\langle X \rangle > x\})$. Therefore, we hope that the results in this paper will be important in better understanding the behavior of the supremum distribution of Gaussian processes with stationary increments, and in analyzing the queue length distribution for heterogeneous types of network traffic.

References

[1] Addie, R. G. and Zukerman, M. (1994). An Approximation for Performance Evaluation of Stationary Single Server Queues. *IEEE Transactions on Communications* **42,** 3150–3160.

[2] Asmusssen, S. (1992). On Cycle Maxima, First Passage Problems and Extreme Value Theory for Queues. *Stochastic Models* **8,** 163–179.

[3] Beran, J., Sherman, R., Taqqu, M. S. and Willinger, W. (1995). Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications* **43,** 1566–1579.

[4] Botvich, D. D. and Duffield, N. G. (1995). Large deviations, the shape of the loss curve, and economies of scale in large multiplexers. *Queueing Systems* **20,** 293–320.

[5] Choe, J. and Shroff, N. B. (1998). A Central Limit Theorem Based Approach for Analyzing Queue Behavior in High-Speed Networks. *IEEE/ACM Transactions on Networking* **6,** 659–671.

[6] Choe, J. and Shroff, N. B. (1999). On the supremum distribution of integrated stationary Gaussian processes with negative linear drift. *Advances in Applied Probability* **31,** 135–157.

[7] Choe, J. and Shroff, N. B. (1999). Queueing analysis with gaussian inputs including srd, lrd, and self-similar processes. *Technical report* TR-ECE 99-12. School of Electrical and Computer Engineering, Purdue University West Lafayette, Indiana.

[8] Duffield, N. G. and O'Connell, N. (1995). Large deviations and overflow probabilities for the general single server queue, with application. *Proc. Cambridge Philos. Soc.* **118,** 363–374.

[9] Feller, W. (1968). *An Introduction to Probability Theory and its Applications I.* John Wiley & Son, New York.

[10] Glynn, P. W. and Whitt, W. (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *Journal of Applied Probability* 131–155.

[11] Leland, W. E., Taqqu, M., Willinger, W. and Wilson, D. V. (1994). On the Self-Similar Nature of Ethernet Traffic (Extended Version). *IEEE/ACM Transactions on Networking* **2,** 1–15.

[12] Loynes, R. M. (1962). The Stability of a Queue with Non-independent Inter-arrival and Service Times. *Proc. Cambridge Philos. Soc.* **58,** 497–520.

[13] M. Montgomery and G. De Veciana (1996). On the Relevance of Time Scales in Performance Oriented Traffic Characterization. In *Proceedings of IEEE INFOCOM.* San Francisco, CA. pp. 513–520.

[14] Maglaris, B., Anastassiou, D., Sen, P., Karlsson, G. and Robbins, J. D. (1988). Performance Models of Statistical Multiplexing in Packet Video Communication. *IEEE Transactions on Communications* **36,** 834–843.

[15] Narayan, O. (1998). Exact Asymptotic Queue-Length Distribution for Fractional Brownian Traffic. *Adv. Perf. Analysis* **1**.

[16] Norros, I. (1995). On the Use of Fractal Brownian Motion in the Theory of Connectionless Networks. *IEEE Journal on Selected Areas in Communications* **13,** 953–962.

[17] Piterbarg, V. I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields.* American Mathematical Society, Providence, RI.

[18] Shroff, N. B. and Schwartz, M. (1998). Improved Loss Calculations at an ATM Multiplexer. *IEEE/ACM Transactions on Networking* **6,** 411–422.

[19] Takacs, L. (1967). *Combinatorial Methods in the Theory of Stochastic Processes.* John Wiley & Son, New York.