UNIFORM RENEWAL THEORY WITH APPLICATIONS TO EXPANSIONS OF RANDOM GEOMETRIC SUMS

J. BLANCHET,* *Harvard University*

P. GLYNN,** *Stanford University*

Abstract

Consider a sequence $X = (X_n : n \geq 1)$ of independent and identically distributed random variables, and an independent geometrically distributed random variable M with parameter p. The random variable $S_M = X_1 + \cdots + X_M$ is called a geometric sum. In this paper we obtain asymptotic expansions for the distribution of S_M as $p \searrow 0$. If $\mathrm{E}\, X_1 > 0$, the asymptotic expansion is developed in powers of p and it provides higher-order correction terms to Renyi's theorem, which states that $\mathrm{P}(pS_M > x) \approx \exp(-x / \mathrm{E}\, X_1)$. Conversely, if $\mathrm{E}\, X_1 = 0$ then the expansion is given in powers of $\sqrt{p}$. We apply the results to obtain corrected diffusion approximations for the M/G/1 queue. These expansions follow in a unified way as a consequence of new uniform renewal theory results that are also developed in this paper.

Keywords: Renewal theory; random geometric sum; corrected diffusion approximation

2000 Mathematics Subject Classification: Primary 60K05; 60K25
Secondary 90B22; 90B20

1. Introduction

Consider a sequence $X = (X_k : k \geq 1)$ of independent and identically distributed (i.i.d.) random variables (RVs), and let M be an independent geometrically distributed RV with mass function

$$\mathrm{P}(M = k) = p(1 - p)^{k-1} = pq^{k-1}, \qquad k \geq 1.$$

Renyi's theorem for geometric sums of random variables establishes that if $0 < \mathrm{E}\, X_1 < \infty$ then

$$\mathrm{P}(pS_M > x) = \exp\left(\frac{-x}{\mathrm{E}X_1}\right) + o(1) \quad \text{as } p \searrow 0, \tag{1}$$

where $S_M = X_1 + \cdots + X_M$ is a geometric sum. One of the main objectives of this paper is to provide (under regularity conditions to be discussed in Sections 2 and 3) higher-order correction terms (in powers of p) to approximation (1).

The asymptotic expansions that we will develop are closely related to the so-called 'corrected diffusion approximations' (CDAs). CDAs for random walks were introduced by Siegmund (1979), who provided a correction for the Brownian approximation to the distribution of the maximum of a random walk (RW) with small negative drift. In order to see the connection

to CDAs, note that Skorohod's embedding (see, for instance, Durrett (2005)) shows that if $\sigma^2 = \text{var}(X_1) < \infty$ then

$$S_{\lfloor t \rfloor} = (\text{E}\,X_1)t + \sigma B(t) + o(t^{1/2}) \quad \text{almost surely (a.s.) as } t \nearrow \infty, \tag{2}$$

where $B(\cdot)$ is a standard Brownian motion and, if $t \geq 0$, $\lfloor t \rfloor$ denotes the integer part of t. As a consequence, (1) follows directly from (2) by letting $t = M$ and using the fact that $pM \overset{w}{\to} \text{Exp}(1)$ as $p \searrow 0$, where $\text{Exp}(1)$ is an exponential RV with unit mean and '$\overset{w}{\to}$' denotes weak convergence. Note, in addition, that when $\text{E}\,X_1 = 0$, (2) also implies that

$$p^{1/2} S_M \overset{w}{\to} \frac{\sigma B(\text{Exp}(1))}{2^{1/2}} \quad \text{as } p \searrow 0, \tag{3}$$

where $B(\text{Exp}(1))$ follows the double exponential or Laplace's distribution (see Kalashnikov (1997)). Here we also develop higher-order correction terms that complement the weak convergence result in (3). As we shall see, in the zero mean case, the asymptotic expansion developed is obtained in powers of $p^{1/2}$, which may have been expected given the scaling present in (3).

We obtain our asymptotic expansions (for the case of positive mean and for the case of zero mean) via a unified approach. In particular, the asymptotic expansions developed here follow as a consequence of a new uniform renewal theorem that is of independent interest. This uniform renewal theorem is closely related to previous such results established by, for example, Siegmund (1979), Borovkov and Foss (2000), and Fuh (2004). A crucial difference between our uniform renewal theorem and previous results is that we obtain uniformity even over families of increment distributions in which the increment's mean can be arbitrarily close to 0. This feature is required in order to obtain the asymptotic expansions in the zero mean case corresponding to (3).

As we have discussed, our theory essentially develops CDAs for geometric sums. However, the theory presented here has the simplifying characteristic that the form of the expansion can be given in terms of moments of the X_is, whereas the CDA for the maximum of a RW has coefficients that are traditionally expressed in terms of one-dimensional integrals involving the increment distribution's characteristic function (rather than in terms of moments of the ladder height RV).

The RV S_M arises in many applied probability settings, as made clear in the book on geometric random sums by Kalashnikov (1997). In particular, the maximum of a RW can be represented as a geometric sum of ladder height RVs. It is well known that the maximum of a RW is of central importance in queueing theory, where it describes the steady-state waiting-time distribution for the single-server queue. Tail probabilities for the maximum of a RW also play a key role in computing infinite horizon ruin probabilities in the insurance setting, as well as boundary crossing probabilities for certain one-sided sequential statistical tests; see Asmussen (2003), Asmussen (2001), and Siegmund (1985). Letting $p \searrow 0$ is natural in the problem settings just described, as this corresponds to 'heavy traffic' in the queueing context and to the 'low safety loading' regime in the insurance environment.

As noted above, the theory developed here establishes an expansion that can be computed in terms of moments of X_1. Hence, in order to apply the results of this paper to the maxima of a RW, we need to be able to express the moment of the ladder height RVs and p as a power series in the natural problem parameterization that arises in the RW context. This computation is straightforward in the setting of the so-called M/G/1 queue, but is in general difficult. Siegmund (1979) calculated the first three terms in the power series for RWs having exponential moments,

whereas Chang and Peres (1997) obtained the entire power series for Gaussian increments. In Blanchet and Glynn (2006) the entire power series was developed for general non-Gaussian increment distributions having exponential moments and strongly nonlattice distributions.

Another important setting in which geometric random sums arise is in the study of the total 'reward' accumulated up to hitting a set in the regenerative Markov process setting (a special case of which is the time required to hit the set). The asymptotic regime of this paper then corresponds to the case in which we study the sequence of sets that become progressively 'rarer' in this context; the basic exponential limit law, (1), is developed in, for example, Keilson (1979) and Aldous (1989). So our current theory offers the opportunity to develop additional correction terms for approximating such distributions. Additional applications of geometric sums (to program debugging and reliability modeling) are discussed in Kalashnikov (1997).

The main contributions of this paper are as follows.

1. We obtain a uniform version of reminder term estimates in the renewal theorem owing to Stone (1965) and Carlsson (1983) over a family of distributions having increment means arbitrarily close to 0 (Theorem 1).

2. When $E X_1 > 0$ is positive and the increments are strongly nonlattice, we develop asymptotic expansions for the distribution of $p S_M$ in powers of p (Theorem 2).

3. When $E X_1 = 0$ (also under a strong nonlattice condition), we provide an asymptotic expansion for the distribution of the RV $p^{1/2} S_M$ in powers of $p^{1/2}$ (Theorem 6).

4. We generalize Theorem 2 to cover the case in which the distribution of the increment X_1 is itself allowed to depend on p (Theorem 4). The generalization is required in order that the theory here can be applied to the maxima of a RW. In Blanchet and Glynn (2007) the authors show how Theorem 4 leads to a CDA for heavy-tailed random walks in which the number of terms in the asymptotic expansion depends on the number of moments assumed to be finite, thereby generalizing the CDA of Hogan (1986) (in which the first two terms of the heavy-tailed expansion were obtained).

5. We develop a CDA for the distribution of the steady-state waiting time for the M/G/1 queue, for both light-tailed and heavy-tailed service-time distributions, generalizing the results of Asmussen and Binswanger (1997) and Abate *et al*. (1995).

This paper is organized as follows. In Section 2 we introduce the uniform renewal theorem that plays a central role in our subsequent analysis. In Section 3 the asymptotic expansions for $P(p S_M > x)$ are derived for the case in which the X_is are nonnegative and possess a finite moment generating function. In Section 4 we develop the asymptotic expansions for nonnegative increments under the existence of at least a finite second moment. In Section 5 we consider cases in which the distribution of the X_is also varies with p. Our expansions are extended to the case of real-valued increments with positive mean in Section 6. In Section 7 we obtain the expansion for zero-mean increment distributions. In Section 8 we apply our results to develop CDAs for the M/G/1 queue.

2. A uniform renewal theorem

In this section we develop the necessary renewal theory that will be key to the rest of this paper. In particular, we will study the renewal function corresponding to a distribution F having support on the entire real line (i.e. F need not be the distribution of a nonnegative RV). Let τ be a RV having distribution F. (When we apply the results of this section to geometric random

sums in later sections, the increment RV X_1 will play the role of τ.) For appropriately integrable g, let $E_F\, g(\tau)$ be defined as

$$E_F\, g(\tau) := \int_{-\infty}^{\infty} g(t) F(\mathrm{d}t),$$

and let us write $\mu_F = E_F\, \tau$. Let us set

$$U_F(t) := \sum_{n=0}^{\infty} F^{*n}(t),$$

and define $H_1^F(\cdot)$ and $H_2^F(\cdot)$ as

$$H_1^F(t) = \begin{cases} \displaystyle\int_t^{\infty} (1 - F(s))\,\mathrm{d}s, & t \geq 0, \\[2ex] \displaystyle\int_{-\infty}^{t} F(s)\,\mathrm{d}s, & t < 0, \end{cases}$$

and

$$H_2^F(t) = \begin{cases} \displaystyle -\int_t^{\infty} H_1^F(s)\,\mathrm{d}s, & t \geq 0, \\[2ex] \displaystyle\int_{-\infty}^{t} H_1^F(s)\,\mathrm{d}s, & t < 0. \end{cases}$$

A distribution F is said to be *strongly nonlattice* if there exists $\varepsilon > 0$ such that

$$\inf_{|\lambda| > \varepsilon} |1 - \chi_F(\lambda)| > 0,$$

where $\chi_F(\lambda) = E_F\, \mathrm{e}^{\mathrm{i}\lambda\tau}$. A family $\mathcal{F}$ of distribution functions is said to be *uniformly strongly nonlattice* if

$$\inf_{F \in \mathcal{F}} \inf_{|\lambda| > \varepsilon} |1 - \chi_F(\lambda)| > 0. \tag{4}$$

The first result of this section is the uniform renewal theorem that will be essential to developing the asymptotic expansions in this paper.

Theorem 1. *Suppose that $\mathcal{F}$ is strongly nonlattice with $E_F\, \tau > 0$ for each $F \in \mathcal{F}$. Then we have the following.*

(i) *If $\sup_{F \in \mathcal{F}} E_F\, \mathrm{e}^{\eta|\tau|} < \infty$ for some $\eta > 0$ then we can find positive constants K_1 and c (independent of $F \in \mathcal{F}$) such that*

$$\mu_F^2 \left| U_F(t) - \left(\frac{t}{\mu_F} + \frac{E_F\, \tau^2}{2\mu_F^2} \right) \mathbf{1}_{[0,\infty)}(t) \right| \leq K_1 \mathrm{e}^{-ct} \quad \text{for } t \geq 0$$

and

$$\mu_F^2 U_F(t) \leq K_1 \mathrm{e}^{ct} + \mu_F K_1 \exp(c\mu_F t) \quad \text{for } t < 0.$$

(ii) *Moreover, under the condition* $\sup_{F \in \mathcal{F}} \mathrm{E}_F \, |\tau|^p < \infty$ *for* $p \geq 3$, *we obtain*

$$\sup_{F \in \mathcal{F}} \mu_F^4 |U_F(t) - \widetilde{U}_F(t)| \leq o(|t|^{-\lfloor p \rfloor} \log(|t|)) \quad as \ |t| \to \infty,$$

where $\widetilde{U}_F(t)$ *satisfies*

$$\widetilde{U}_F(t) = \begin{cases} \dfrac{t}{\mu_F} + \dfrac{\mathrm{E}_F \, \tau^2}{2\mu_F^2} + \dfrac{H_2^F(t)}{\mu_F^2} + \dfrac{(H_1^F * H_1^F)(t)}{\mu_F^3}, & t \geq 0, \\[3mm] \dfrac{H_2^F(t)}{\mu_F^2} + \dfrac{(H_1^F * H_1^F)(t)}{\mu_F^3}, & t < 0. \end{cases} \tag{5}$$

The proof of this result is given at the end of this section. Note that part (i) describes an error term of order μ_F^{-2} whereas part (ii) describes an error term of order μ_F^{-4}. Since part (i) is a special case of part (ii), this suggests that the error term in part (ii) is not the best possible result (as a function of μ_F). The proof technique is based on a Fourier analysis argument due to Carlsson (1983). Our expression for $\widetilde{U}_F(\cdot)$ coincides with the one provided by Carlsson (1983) and, as we shall see in the proof, our argument requires a normalization of the $\widetilde{U}_F$s Fourier transform by the factor μ_F^4 in order to be amenable to uniform estimates over the family $\mathcal{F}$.

Such uniform estimates are the most important part of our contribution here. Uniform renewal theorems have been studied previously in the literature. Siegmund (1979) provided a result similar to part (i). However, Siegmund's result only covered the case of positive increments and the underlying family of distributions enjoyed a particular parameterization through exponential families. The fact that Theorem 1 is developed on the whole line with an explicit dependence on the mean of the increment distributions underlying the renewal function (i.e. μ_F) will be crucial in our developments, especially for the case in which $\mathrm{E} \, X_1 = 0$. Fuh (2004) studied uniform Markov renewal theory, which included uniform renewal theory as a special case, on the whole line under the same assumptions as given in Theorem 1. However, there is no dependence on the mean of the increment distributions of the renewal function in his estimates. The information given in Theorem 1, particularly with respect to μ_F, is key for our purposes because we are interested in understanding how the distribution of S_M changes in general—even when we have a zero-mean increment distribution. Borovkov and Foss (2000) also developed a uniform renewal theorem under nonlattice assumptions, like those imposed here. Again, they considered the case in which the mean of increment distributions was bounded away from 0 and, therefore, that theory does not directly apply to our current situation.

2.1. Proof of Theorem 1

For $F \in \mathcal{F}$, set

$$V_F(x, h, a) = \mathrm{E}[U_F(x + h - aZ) - U_F(x - aZ)],$$

where $Z \sim N(0, 1)$. Lemma 1, below, corresponds to Equation (8) of Stone (1965).

Lemma 1. *If* F *is strongly nonlattice and* $\mu_F = \mathrm{E}_F(\tau) > 0$ *then*

$$V(x, h, a) = \frac{h}{2\mu} + \frac{h}{2\pi \int_{-\infty}^{\infty} \mathrm{Re}[\mathrm{e}^{-\mathrm{i}x\lambda}(1 - \mathrm{e}^{-\mathrm{i}h\lambda})\mathrm{e}^{-a^2\lambda^2/2}/(\mathrm{i}h\lambda(1 - \chi_F(\lambda)))]\,\mathrm{d}\lambda}.$$

Lemma 2, below, is due to Carlsson (1983); see Lemma 1 therein. (See also Theorem 7.3 of Ganelius (1971).)

Lemma 2. *Define* $\mathcal{M}_b$ *to be the class of those nondecreasing functions* $f : \mathbb{R} \to [1, \infty)$ *that satisfy* $f(2x) \le bf(x)$ *and* $f(x) = 1$ *if* $x \le 1$. *Let* $g(\cdot)$ *have Fourier transform given by*

$$\widehat{g}(\theta) = \int_{-\infty}^{\infty} e^{-i\theta s} \, dg(s).$$

Suppose that there exist $K > 0$ *and* $m_T \in \mathcal{M}_b$ *for which the inequality*

$$\sup_{x \le y \le x + 1/T} (g(y) - g(x)) \le \frac{K}{m_T(x)}$$

holds for all $x \in \mathbb{R}$. *If* $(\widehat{g}(\theta) \colon \theta \in \mathbb{R})$ *has its support outside* $(-T, T)$ *then there exists a constant* C *depending only on* b *(independent of* T*) and such that* $|g(x)| \le CK/m_T(x)$.

As we shall see, some of the remainder terms in our estimates satisfy the conditions of Lemma 2. We shall also need the following uniform variant of the Riemann–Lebesgue lemma.

Lemma 3. *Let* $\mathcal{G} = \{g_\alpha(\cdot) \colon \alpha \in \chi\}$ *be a family of real-valued functions defined on the compact interval* $[0, 1]$ *and indexed by an arbitrary set* χ. *Assume that the* $g_\alpha(\cdot)s$ *are uniformly bounded and equicontinuous. Then,*

$$\lim_{n \to \infty} \sup_{\alpha \in \mathcal{I}} \int_0^1 g_\alpha(\lambda) e^{i\lambda n} \, d\lambda = 0.$$

Proof. Define $g_{\alpha,m}(\cdot)$ as

$$g_{\alpha,m}(\lambda) = \sum_{k=0}^{m-1} g_\alpha \left(\frac{k}{m} \right) \mathbf{1}_{[k/m, (k+1)/m)}(\lambda).$$

We then have

$$\left| \sup_{\alpha \in \mathcal{I}} \int_0^1 g_\alpha(\lambda) e^{i\lambda n} \, d\lambda \right| \le \sup_{\alpha \in \mathcal{I}} \int_0^1 |g_\alpha(\lambda) - g_{\alpha,m}(\lambda)| \, d\lambda$$
$$+ \sum_{k=0}^{m-1} \sup_{\alpha \in \mathcal{I}} \left| g_{\alpha,m} \left(\frac{k}{m} \right) \right| \left| \int_{k/m}^{(k+1)/m} e^{i\lambda n} \, d\lambda \right|.$$

The first integral can be made arbitrarily small in m by virtue of equicontinuity, while the second integral, owing to uniform boundedness, is of order $O(1/n)$. This yields the conclusion of Lemma 3.

Before we provide the proof of Theorem 1, let us briefly comment on some of the methods behind its technical development. The proof of part (i) is adapted from the work of Stone (1965), except at the end where we link the behavior of $V_F(x, h, a)$ to that of $U_F(x)$ as $x \nearrow \infty$. For this latter part, we use the results of Carlsson (1983). Both Stone (1965) and Carlsson (1983) relied on the use of Fourier analysis estimates, but Carlsson applied Fourier transforms to distributions because he worked directly with $U_F(\cdot)$ rather than with an expression such as $V_F(\cdot)$ (which involves increments of $U_F(\cdot)$ convolved with a smooth kernel). Basically, analyzing the increments of $U_F(\cdot)$ removes the need for dealing with the Fourier transform

of an indicator function which gives rise to a delta function. Stone (1965) also provided a way to link the asymptotic behavior of $V_F(\cdot, h, a)$ to that of $U_F(\cdot)$, but Carlsson's techniques are slightly more direct and blend in better with our development in part (ii) of Theorem 1. Stone's techniques are convenient in the case of exponential moments because they make it possible to easily remove the singularities that appear when $E_F \tau$ is arbitrarily close to 0 (as will be indicated precisely in the proof below) using Cauchy's theorem for analytic functions. Carlsson's techniques are convenient in part (ii) because his expressions are especially well suited to a direct Fourier analysis of the terms involved in $\widetilde{U}_F$. The rate of decay in the error obtained by approximating U_F by $\widetilde{U}_F$ directly is obtained using such Fourier expressions (via the use of Lemma 2). Carlsson's expressions are homogeneous in μ_F^{-4}, and the estimates depend only on the tail decay of $\overline{F}$ implied by the existence of moments of order p. As a consequence, our uniform assumptions make Carlsson's analysis easily adaptable to our current situation.

Proof of Theorem 1. We shall first prove part (i) in several steps.

Step 1: Set $\upsilon_F = E_F \tau^2$. The idea is to first use Lemma 1 to estimate $V(x, h, a)$. More precisely, we note that

$$\int_{-\infty}^{\infty} e^{-ix\lambda}(1 - e^{-ih\lambda})(ih\lambda)^{-1} e^{-a^2\lambda^2/2}(1 - \chi_F(\lambda))^{-1}\, d\lambda$$

$$= \int_{-\infty}^{\infty} e^{-ix\lambda}(1 - e^{-ih\lambda})(ih\lambda)^{-1} e^{-a^2\lambda^2/2}\left(-i\mu_F\lambda + \frac{\upsilon_F\lambda^2}{2}\right)^{-1} d\lambda \qquad (6)$$

$$+ \int_{-\infty}^{\infty} e^{-ix\lambda}(1 - e^{-ih\lambda})(ih\lambda)^{-1} e^{-a^2\lambda^2/2}$$

$$\times \left((1 - \chi_F(\lambda))^{-1} - \left(-i\mu_F\lambda + \frac{\upsilon_F\lambda^2}{2}\right)^{-1}\right) d\lambda. \qquad (7)$$

Observe that the integrand corresponding to (7) is analytic on a strip defined by the imaginary axis and a line parallel to this axis (both to the right or the left) that can be taken independent of the choice $F \in \mathcal{F}$ (as a consequence of the uniform strongly nonlattice assumption). This is because we are removing the singularity of $1/(1 - \chi_F(\lambda))$ at 0 by subtracting $(-i\mu_F\lambda + \upsilon_F\lambda^2/2)^{-1}$. This step is different from that of Stone (1965) (we have added the term including υ_F because we are also interested in the case in which μ_F can be arbitrarily small). Cauchy's theorem allows us to rewrite (7) as

$$\int_{-\infty}^{\infty} e^{-(c+i\lambda)x}(1 - e^{-(c+i\lambda)h})(c + i\lambda)^{-1} \exp\left(\frac{a^2(c + i\lambda)^2}{2}\right)$$

$$\times \left((1 - \chi_F(c + i\lambda))^{-1} - \left(-\mu_F(c + i\lambda) + \frac{\upsilon_F(c + i\lambda)^2}{2}\right)^{-1}\right) d\lambda$$

for some $c > 0$ sufficiently small (uniformly over the family $\mathcal{F}$). Note that the previous expression is of order $O(e^{-cx}(1 + |\log(a)|))$ as x tends to ∞ uniformly for h in bounded sets and uniformly over the family $\mathcal{F}$. This follows because the integral in the previous display can be bounded by $C_0 \exp(-cx - a^2\lambda^2/2)$ and the choice of the constant C_0 is independent of F by virtue of the strongly nonlattice assumption (also observe that the strongly nonlattice assumption also guarantees $\inf_{F \in \mathcal{F}} \upsilon_F > 0$). Note that a completely analogous estimate can

be obtained as x tends to $-\infty$. Conversely, the integral in (6) can be rewritten as

$$\int_{-\infty}^{\infty} \frac{e^{-ix\lambda}(1 - e^{-ih\lambda})(ih\lambda)^{-1}e^{-a^2\lambda^2/2}}{-i\mu_F\lambda} \, d\lambda \tag{8}$$

$$+ \int_{-\infty}^{\infty} e^{-ix\lambda}(1 - e^{-ih\lambda})(ih\lambda)^{-1}e^{-a^2\lambda^2/2}$$

$$\times \left(\frac{1}{-i\mu_F\lambda + \upsilon_F\lambda^2/2} - \frac{1}{-i\mu_F\lambda} \right) d\lambda. \tag{9}$$

Let us refer to the integrals in (8) and (9) as J_1 and J_2, respectively. The integral J_1 can be evaluated explicitly (see the expression following Equation (8) of Stone (1965)), yielding the conclusion that

$$\mu_F J_1 = \pm\frac{h}{2} + O(e^{-c|x|}) \quad \text{as } x \longrightarrow \pm\infty$$

uniformly over the family $\mathcal{F}$ and h, a in bounded sets. For the integral J_2, we have

$$\frac{4h\mu_F^2 J_2}{2\pi \upsilon_F} = P\left(\frac{x}{2a} \le N(0, 1) \le \frac{x + h}{2a} \right)$$

$$+ P\left(N(0, 1) > \frac{x + 2b_Fa^2}{a} \right) \exp(\mu_F(\mu_Fa^2 + x))$$

$$- P\left(N(0, 1) > \frac{x + h + 2b_Fa^2}{a} \right) \exp(\mu_F(\mu_Fa^2 + x + h)),$$

where $b_F = 2\mu_F/\upsilon_F$. Consequently, $\mu_F^2 J_2 = O(e^{-cx})$ as $x \to \infty$ uniformly over the family $\mathcal{F}$ and for a in bounded intervals, whereas

$$|\mu_F^2 J_2| \le \mu_F K_1 \exp(x\mu_F) \quad \text{as } x \to -\infty$$

uniformly for a in bounded sets with K_1 independent of F. Consequently,

$$\mu_F^2 V_F(x, h, a) = h\mu_F + O(e^{-cx}(1 + |\log(a)|)) \quad \text{as } x \to \infty,$$

where the term $O(e^{-cx}(1 + |\log(a)|))$ does not depend on the choice of F or h, and similarly

$$\mu_F^2 V_F(x, h, a) \le O(e^{cx}(1 + |\log(a)|)) + \mu_F K_1 \exp(x\mu_F) \quad \text{as } x \to -\infty.$$

Step 2: Next we want to use an argument similar to that given by Stone (1965) to estimate $U_F(x + 1) - U_F(x) - \mu_F^{-1}$. This follows easily by noting that

$$\mu_F^2 V_F(x, 2, 1) = \sum_{n=0}^{\infty} \mu_F^2 P(x \le S_n + Z \le x + 2)$$

$$\ge \sum_{n=0}^{\infty} \mu_F^2 P(x \le S_n + Z \le x + 2; \ Z \in [0, 1])$$

$$\ge \mu_F^2 (U_F(x + 1) - U_F(x)) P(Z \in (0, 1)).$$

Consequently, a finite number $K > 0$ (independent of F) can be chosen such that

$$\mu_F^2 (U_F(x + h) - U_F(x)) \le K$$

uniformly over $|h| \leq 2$. Following Stone (1965), fix $\delta > 0$, and note that if $|y| \leq 2e^{-\delta x}$, we have

$$U_F(x + 1 - e^{-\delta x} - y) - U_F(x + e^{-\delta x} - y) \leq U_F(x + 1) - U_F(x)$$
$$\leq U_F(x + 1 + e^{-\delta x} - y) - U_F(x - e^{-\delta x} - y).$$

Now, choose $x_0 > 0$ such that $P(|N(0, 1)| \geq x_0) \leq \exp(-\delta x_0) \leq \frac{1}{2}$. It follows, as a consequence of the previous estimates, that, for $x \geq x_0$,

$$\mu_F^2 V_F\left(x + e^{-\delta x}, 1 - 2e^{-\delta x}, \frac{e^{-\delta x}}{x}\right) - \mu_F^2 K e^{-\delta x}$$
$$\leq \mu_F^2 (U_F(x + 1) - U_F(x))$$
$$\leq \mu_F^2 (1 - e^{-\delta x})^{-1} V_F\left(x - e^{-\delta x}, 1 + 2e^{-\delta x}, \frac{e^{-\delta x}}{x}\right).$$

This implies that there exists $c > 0$ such that

$$\mu_F^2 (U_F(x + 1) - U_F(x) - \mu_F^{-1}) = O(e^{-cx}) \tag{10}$$

uniformly over $\mathcal{F}$ as x tends to ∞, and

$$\mu_F^2 |U_F(x + 1) - U_F(x)| = O(e^{cx}) + \mu_F K e^{x\mu_F} \tag{11}$$

also uniformly over $\mathcal{F}$ as x tends to $-\infty$.

Step 3: Finally, we proceed to show that

$$\mu_F^2\left(U_F(x) - \left(x\mu_F^{-1} + \frac{\upsilon_F}{2\mu_F}\right)\mathbf{1}_{[0,\infty)}(x)\right) = o(1) \tag{12}$$

uniformly over $\mathcal{F}$ as x tends to ∞. As in Carlsson (1983), define

$$G_F(x) := U_F(x) - \left(x\mu_F^{-1} + \frac{\upsilon_F}{2\mu_F}\right)\mathbf{1}_{[0,\infty)}(x) + \frac{H_2^F(x)}{\mu_F^2}.$$

Carlsson (1983) evaluated the Fourier transform $\widehat{G}_F$ of G_F,

$$\widehat{G}_F(\lambda) := (\chi_F(\lambda) - 1 + i\lambda\mu_F)^2 (i\lambda)^{-1} (i\lambda\mu_F)^{-2} (1 - \chi_F(\lambda))^{-1};$$

see the end of Section 3 of Carlsson (1983). Estimate (12) will follow after applying Lemma 2 to a suitable modification of $G_F(\cdot)$. In particular, the strategy is to consider $G_F^T(\cdot)$ and $G_F^*(\cdot)$, which are defined via

$$G_F^T(x) := G_F(x) - \frac{1}{2\pi} \int_{-T}^{T} \widehat{G}_F(\lambda) e^{i\lambda x} \, d\lambda$$
$$:= G_F(x) - G_F^*(x).$$

Note that $\widehat{G_F^T}(\lambda)$ is 0 for $|\lambda| \leq T$. Conversely, using Chebyshev's inequality, we can find a constant K (independent of F) such that, for $x \geq 1$ and $T \geq 1$,

$$\sup_{x \leq y \leq x + 1/T} \mu_F^2 |H_2^F(x) - H_2^F(y)| \leq \mu_F^2 (H_2^F(x) - H_2^F(x + T^{-1})) \leq \frac{K}{T}.$$

Also, observe that $\mu_F^3 G_F^*(\cdot)$ is bounded (for fixed T) and that

$$\lim_{x \to 0} \sup_F |\mu_F^3 G_F^*(x)| = 0. \tag{13}$$

This last step follows by invoking Lemma 3, which can be safely applied by means of the strongly nonlattice condition, the form of $\widehat{G}_F$, and the existence of uniform bounds in moments of order at least 2. These observations, combined with our previous estimates concerning $U_F(x+h) - U_F(x)$, yield

$$\sup_{x \le y \le x + 1/T} \mu_F^2 (G_F^T(x) - G_F^T(y)) \le \frac{K}{T}$$

for some positive constant $K > 0$ (independent of F). Consequently, Lemma 2, applied to G_F^T, together with (13) yield

$$\lim_{x \to \infty} \sup_F \mu_F^2 |G_F(x)| \le \frac{CK}{T}$$

for some constant C independent of T. Since T was arbitrary, we obtain (12). This combined with (10) and (11) gives part (i).

Part (ii) is similar to the end of part (i) and the proof proceeds by following a similar program as in Carlsson (1983). Again a key assumption is the uniform strongly nonlattice condition, (4). The Fourier inversion expressions are provided by Carlsson (1983) for each fixed F. First we define $(\widehat{R}(\lambda)\colon \lambda \in \mathbb{R})$ via

$$\widehat{R}_F(\lambda) = \mu_F^{-4} \left(\frac{\chi(-\lambda) - 1 + i\lambda \mu_F}{(i\lambda)^2} \right)^2 \left(\frac{\chi(-\lambda) - 1 + i\lambda \mu_F}{i\lambda} \right).$$

Then, it follows that the inverse Fourier transform of $\widehat{R}$ is given by

$$R_F(t) = \mu_F^{-4} (H_1^F \star H_1^F \star \dot{H}_1^F)(t),$$

where $\dot{H}_1^F$ denotes the derivative of H_1^F (the derivative can be interpreted weakly or we can make sense out of $R_F(t)$ using integration by parts by noting that $H_1^F \star H_1^F$ is differentiable). It easily follows that $\mu_F^4 R_F(t) = o(|t|^{-\lfloor p \rfloor})$ as $|t| \nearrow \infty$ uniformly over the class $\mathcal{F}$ assuming that $\sup_{F \in \mathcal{F}} \mathrm{E}_F |\tau|^p < \infty$. Now, the Fourier transform of $\Gamma_F = \widetilde{U}_F - R_F$ (see the definition in part (ii) of the present theorem) is $\widehat{\Gamma}_F(\cdot)$ given by

$$\mu_F^4 \widehat{\Gamma}_F(\lambda) = (\chi_F(-\lambda) - 1 + i\lambda \mu_F)^4 (i\lambda)^{-5} (1 - \chi_F(-\lambda)).$$

The estimates for Γ_F again involve applications of Lemma 2 following Proposition 1 of Carlsson (1983) and the use of Lemma 3 just as in the case of G_F in part (i), above.

Condition (4) is crucial in order to guarantee the validity of our uniform renewal results. It will often be useful and more convenient to verify the following alternative version of (4).

Lemma 4. *Suppose that $\mathcal{F}$ is a tight family of distribution functions satisfying*

$$\inf_{F \in \mathcal{F}} \inf_{|\lambda| \ge 1} |1 - \chi_F(\lambda)| > 0. \tag{14}$$

Then, the family $\mathcal{F}$ is uniformly strongly nonlattice in the sense that (4) is satisfied.

Remark 1. Note that if $\sup_{F \in \mathcal{F}} \mathrm{E}_F |\tau| < \infty$ then the family $\mathcal{F}$ is tight. In order to apply Theorem 1, we must verify the moment assumptions indicated in its parts (i) or (ii). Hence, it suffices to just verify condition (14) instead of (4).

Proof of Lemma 4. Fix $\varepsilon > 0$, and note that

$$\inf_{F \in \mathcal{F}} \inf_{|\lambda| \geq \varepsilon} |1 - \chi_F(\lambda)| = \inf_{F \in \mathcal{F}} \min\Big(\inf_{|\lambda| \geq 1} |1 - \chi_F(\lambda)|, \inf_{\varepsilon \leq |\lambda| \leq 1} |1 - \chi_F(\lambda)| \Big),$$

so we just have to show that

$$\gamma_F = \inf_{\varepsilon \leq |\lambda| \leq 1} |1 - \chi_F(\lambda)|$$

satisfies $\inf_{F \in \mathcal{F}} \gamma_F > 0$. Suppose that this is not the case, then there exists a sequence of distributions $(F_n : n \geq 1)$ in $\mathcal{F}$ and a sequence $(\lambda_n : n \geq 1)$ (with $|\lambda_n| \in [\varepsilon, 1]$) such that $\chi_{F_n}(\lambda_n) \to 1$ as $n \to \infty$. Since $\mathcal{F}$ is tight, there must be a subsequence of distributions $(F_{n_k} : k \geq 1)$ such that $F_{n_k} \xrightarrow{\mathrm{w}} G$ (G may not necessarily be in $\mathcal{F}$) and a subsequence $(\lambda_{n_k} : k \geq 1)$ such that $\lambda_{n_k} \to \lambda$ with $|\lambda| \in [\varepsilon, 1]$, and $\chi_G(\lambda) = 1$, this implies that G is lattice. On the other hand, we have

$$0 < \lim_{k \to \infty} |1 - \chi_{F_{n_k}}(\lambda)| = |1 - \chi_G(\lambda)|,$$

which implies that G is nonlattice, yielding a contradiction and proving the result.

3. Nonnegative geometric sums with exponential moments

Our results in this section improve upon Renyi's approximation, (1), assuming the existence of exponential moments. Define $\phi(\eta) := \mathrm{E} \exp(\eta X_1) < \infty$ and $\phi'(\eta) := \mathrm{d}\phi(\eta)/\mathrm{d}\eta$.

Theorem 2. *Suppose that X_1 has a strongly nonlattice distribution and that $\phi(\eta) < \infty$ for some $\eta > 0$. Then, there exist constants $a, \kappa > 0$ such that, for sufficiently small p,*

$$\left| \mathrm{P}(p S_M > x) - \exp\Big(-\frac{x\tilde{\theta}}{p} + r(p) \Big) \right| \leq \kappa \exp\Big(-\frac{ax}{p} \Big) \tag{15}$$

uniformly in $x \geq 0$, where $\tilde{\theta}$ is the unique nonnegative solution to $\mathrm{E} \exp(\tilde{\theta} X_1) = q^{-1}$ (which exists if $p > 0$ is small enough), and

$$\mathrm{e}^{r(p)} = \frac{p}{q^2 \tilde{\theta} \phi'(\tilde{\theta})} := c(p). \tag{16}$$

Remark 2. We state our approximation in terms of $r(p)$ because later we shall be interested in suggesting approximations that are given as asymptotic expansion powers of p. Using the asymptotic expansions for $r(\cdot)$ instead of $c(\cdot)$ preserves the positivity of our suggested approximation.

Theorem 2 provides rigorous support for the approximation

$$\mathrm{P}(S_M > y) \approx \exp(-y\tilde{\theta} + r(p)). \tag{17}$$

Note that, for a fixed value p, the evaluation of $\tilde{\theta}$ and $r(p)$ (or equivalently, $c(p)$) is a straightforward numerical task and, thus, (17) can be easily implemented to provide an accurate approximation for the distribution of S_M that can be expected to improve upon Renyi's approximation.

Once we know that (17) holds uniformly for small values of p, then we can easily provide (via the implicit function theorem) an expansion for $\tilde{\theta}$ in powers of p and, as a consequence, for $r(p)$. We record these observations in Proposition 1, below.

Proposition 1. *Let $\tilde{\theta}$ and r be defined as in Theorem 2. Then $\tilde{\theta}(p)$ and $r(p)$ can be expanded in absolutely convergent power series in $p \in [0, \delta]$ for some $\delta > 0$.*

Proof. The fact that the series converges absolutely just follows by applying the inverse function theorem (for complex-valued functions) to the analytic extension of $\phi(\cdot)$ at 0.

Remark 3. If our interest is to provide accurate approximations for the distribution of S_M for many different small values of p, an asymptotic expansion such as that suggested by Theorem 2 and Proposition 1 for $P(S_M > \cdot)$ will typically provide a more convenient approximation (in terms of the computational cost) than (17). Let us then provide the elements of the expansion described in Proposition 1, thereby giving the expansion of $P(S_M > x/p)$ in powers of p. Proposition 1 establishes that

$$\tilde{\theta}(p) = \sum_{k=1}^{\infty} \frac{\tilde{\theta}^{(k)}(0) p^k}{k!} \quad \text{and} \quad r(p) = \sum_{k=0}^{\infty} \frac{r^{(k)}(0) p^k}{k!}.$$

For notational convenience, let us write $\tilde{\theta}^{(k)}(0)/k! = \gamma_k$ and $r^{(k)}(0)/k! = \xi_k$. We know that $\tilde{\theta}^{(1)}(0) = \gamma_1 = 1/E X_1$ and $r(0) = 0$, the rest of the γ_ks and ξ_ks can be easily computed via the implicit function theorem. For instance, $2 E^3 X_1 \gamma_2 = 2 E^2 X_1 - E X_1^2$, $6 E^5 X_1 \gamma_3 = 3 E X_1^2 - 6 E X_1^2 E^2 X_1 + 6 E^4 X_1 - E X_1 E X_1^3$. Consequently, $2 E^2 X_1 \xi_1 = 2 E^2 X_1 + E X_1^2$ and $24 E^4 X_1 \xi_2 = 12 E^4 X_1 + 12 E X_1^2 E^2 X_1 - 9 E^2 X_1^2 + 4 E X_1^3 E X_1$. A similar expansion for $\tilde{\theta}(\cdot)$ has been obtained by Abate *et al.* (1995) in connection with heavy-traffic asymptotics for queues. We shall discuss these types of applications in Section 8, below.

For completeness we provide a set of recursive equations to compute the γ_ks.

Proposition 2. *For $n \geq 1$ and each $k \leq n$, the constants $(\gamma_k : 1 \leq k \leq n)$ can be computed by solving recursively the following set of equations (note that the kth equation is linear in γ_k and it depends only on the γ_js for $j \leq k$):*

$$\sum_{m=1}^{k} \frac{E X_1^m}{m!} \sum_{\{n_1 + \cdots + n_m = k-m, n_1, \ldots, n_m \geq 0\}} \prod_{j=1}^{m} \gamma_{n_j+1} = 1 \quad \text{for } 1 \leq k \leq n.$$

Proof. The proof follows directly by applying the implicit function theorem. The details are omitted.

Theorem 2 therefore provides the means to develop an algorithm, which can be implemented easily, for computing an asymptotic expansion for the tail probability $P(S_M > x/p)$ in powers of p that gives the desired corrected diffusion approximation in the present geometric sum setting.

We now discuss the mathematical development of Theorem 2. Let $a(t) = P(S_M > t)$, and note that

$$a(t) = P(X_1 > t) + q \int_{[0,t)} a(t - s) P(X_1 \in \mathrm{d}s). \tag{18}$$

If $p > 0$ is small enough and $E \exp(\eta X_1) < \infty$ for some $\eta > 0$ then the equation

$$\phi(\tilde{\theta}) = E \exp(\tilde{\theta} X_1) = \frac{1}{q} \tag{19}$$

has a unique solution $\tilde{\theta} > 0$. Therefore, making use of well-known techniques, (18) can be transformed into the nondefective renewal equation

$$\exp(\tilde{\theta}t)a(t) = \exp(\tilde{\theta}t)\,\mathrm{P}(X_1 > t) + \int_{[0,t)} \exp(\tilde{\theta}(t-s))a(t-s)F_{\tilde{\theta}}(\mathrm{d}s),$$

where $F_{\tilde{\theta}}(\mathrm{d}s) = q\exp(\tilde{\theta}s)\,\mathrm{P}(X_1 \in \mathrm{d}s)$. Renewal theory then implies that

$$\exp(\tilde{\theta}t)a(t) = \int_{[0,t)} \exp(\tilde{\theta}(t-s))\,\mathrm{P}(X_1 > t-s)U_{\tilde{\theta}}(\mathrm{d}s), \qquad (20)$$

where $U_{\tilde{\theta}}(t) = \mathrm{E}_{\tilde{\theta}}(N(t)+1)$, and under $\mathrm{P}_{\tilde{\theta}}$ the X_is are i.i.d. with distribution $F_{\tilde{\theta}}$. Using the results of Stone (1965), it is not hard to verify that, for fixed but small $p > 0$,

$$\left| \exp(\tilde{\theta}t)a(t) - \frac{1}{\mathrm{E}_{\tilde{\theta}}X_1} \int_{[0,t)} \exp(\tilde{\theta}s)\,\mathrm{P}(X_1 > s)\,\mathrm{d}s \right| \le K(p)\mathrm{e}^{-a(p)t}. \qquad (21)$$

Note that Stone's estimates provide exponential rates of convergence for fixed $p > 0$, but they do not say anything about the behavior of $K(p)$ and $a(p)$ in (21) for small values of p. In contrast, as the proof that follows next indicates, Theorem 1 allows us to set $t = x/p$ and obtain an exponential rate of convergence, thereby controlling the behavior of $K(p)$ and $a(p)$ as $p \searrow 0$.

Proof of Theorem 2. The previous argument leads us to (20). We now verify that the assumptions in Theorem 1 are satisfied. Let us define $g_{\tilde{\theta}}(\lambda) := \mathrm{E}_{\tilde{\theta}}\exp(\mathrm{i}\lambda X_1) = q\,\mathrm{E}\exp((\mathrm{i}\lambda + \tilde{\theta})X_1)$. Using the implicit function theorem on (19), it easily follows that $\tilde{\theta} = p/\mathrm{E}\,X_1 + O(p^2)$. As a consequence, the following inequality can be easily derived for all $p > 0$ sufficiently small and some $M_1 \in (0, \infty)$,

$$|g_{\tilde{\theta}}(\lambda) - g(\lambda)| \le M_1 p.$$

Hence, we conclude that it is possible to choose $\delta > 0$ sufficiently small so that

$$\inf_{p \in [0,\delta]} \inf_{|\lambda| \ge 1} |1 - g_{\tilde{\theta}}(\lambda)| \ge \inf_{p \in [0,\delta]} \inf_{|\lambda| \ge 1} |1 - g(\lambda)| - M_1\delta > 0,$$

which verifies condition (14) in Lemma 4. Finally, because $\tilde{\theta} = O(p)$, it is possible to choose $p > 0$ small enough so that $\mathrm{E}_{\tilde{\theta}}\exp(\eta X_1) = q\,\mathrm{E}_{\tilde{\theta}}\exp((\eta + \tilde{\theta})X_1) < \infty$ for some $\eta > 0$. We can now apply Theorem 1 to (20) and obtain

$$\left| \exp\left(\tilde{\theta}\frac{x}{p}\right)a\left(\frac{x}{p}\right) - \frac{1}{\mathrm{E}_{\tilde{\theta}}X_1} \int_0^\infty \exp(\tilde{\theta}s)\,\mathrm{P}(X_1 > s)\,\mathrm{d}s \right|$$

$$\le \frac{1}{\mathrm{E}_{\tilde{\theta}}X_1} \int_{x/p}^\infty \exp(\tilde{\theta}s)\,\mathrm{P}(X_1 > s)\,\mathrm{d}s \qquad (22)$$

$$+ \left| \frac{1}{\mathrm{E}_{\tilde{\theta}}X_1} \int_{[0,x/p)} \exp\left(\tilde{\theta}\left(\frac{x}{p}-s\right)\right)\mathrm{P}\left(X_1 > \frac{x}{p}-s\right)V(\mathrm{d}s) \right|, \qquad (23)$$

where $V(\cdot)$ is a function that we are introducing here and it corresponds to the left-hand side of the first equation in part (i) of Theorem 1, therefore $|V(t)| = O(\mathrm{e}^{-at})$ for some $a > 0$. The integral in (22) is easily seen to be bounded by $K\mathrm{e}^{-ax/p}$ for some finite constants $K, a > 0$

(assuming that $p > 0$ is sufficiently small). We just need to analyze the integral in (23). Integration by parts yields

$$\int_{[0,x/p)} \exp\left(\tilde{\theta}\left(\frac{x}{p} - s\right)\right) P\left(X_1 > \frac{x}{p} - s\right) V(ds)$$

$$= V\left(\frac{x}{p}\right) P(X_1 > 0) - \exp\left(\tilde{\theta}\frac{x}{p}\right) P\left(X_1 > \frac{x}{p}\right) V(0) \tag{24}$$

$$+ \tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) \int_{[0,x/p)} V(s) \exp(-\tilde{\theta}s) P\left(X_1 > \frac{x}{p} - s\right) ds \tag{25}$$

$$+ \exp\left(\tilde{\theta}\frac{x}{p}\right) \int_{[0,x/p)} V(s) \exp(-\tilde{\theta}s) P\left(X_1 > \frac{x}{p} - ds\right). \tag{26}$$

The absolute value of (24) is also bounded by $K\mathrm{e}^{-ax/p}$ for some finite constants $K, a > 0$. For the integral in (25), observe that

$$\tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) \left| \int_{[0,x/p)} V(s) \exp(-\tilde{\theta}s) P\left(X_1 > \frac{x}{p} - s\right) ds \right|$$

$$\leq \tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) \int_{[0,x/2p)} |V(s)| \exp(-\tilde{\theta}s) P\left(X_1 > \frac{x}{p} - s\right) ds$$

$$+ \tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) \int_{[x/2p,x/p)} |V(s)| \exp(-\tilde{\theta}s) P\left(X_1 > \frac{x}{p} - s\right) ds$$

$$\leq \tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) P\left(X_1 > \frac{x}{2p}\right) M + \tilde{\theta} \exp\left(\tilde{\theta}\frac{x}{p}\right) \int_{[x/2p,\infty)} |V(s)| ds.$$

Since $\tilde{\theta} = O(p)$ and X_1 has exponential moments, we conclude that the previous expression is bounded by $K\mathrm{e}^{-ax/p}$ (for appropriate positive constants K and a). The treatment for the integral in (26) is very similar to that of (25). Thus, we conclude that

$$\mathrm{E}\, q^{N(x/p)} = \frac{\int_0^\infty \exp(\tilde{\theta}s) P(X_1 > s)\, ds}{\mathrm{E}_{\tilde{\theta}} X_1} + O(\mathrm{e}^{-ax/p}). \tag{27}$$

In order to recover the required expression for $c(p)$, note that

$$\mathrm{E}_{\tilde{\theta}} X_1 = q \int_{[0,\infty)} s \exp(\tilde{\theta}s) P(X_1 \in ds) = q\phi'(\tilde{\theta}).$$

Conversely, using integration by parts and the definition of $\tilde{\theta}$, we see that

$$\int_0^\infty \exp(\tilde{\theta}s) P(X_1 > s)\, ds = \frac{\phi(\tilde{\theta}) - 1}{\tilde{\theta}} = \frac{p}{q\tilde{\theta}}.$$

Combining the previous two identities together in (27) yields (15). The analytic properties of $\tilde{\theta}$ follow directly from the implicit function theorem. It is easy to see that $r(\cdot)$ is well defined at 0 (i.e. that the right-hand side of (16) is strictly positive when p is close to 0). However, it is almost immediate to verify that $c(\cdot)$ is real analytic at the origin with $c(0) = 1$. This implies the real analyticity of r and the conclusion of the theorem.

4. Nonnegative geometric sums with heavy tails

In this section we also assume that $X = (X_n : n \geq 1)$ is an i.i.d. sequence of nonnegative RVs and, in contrast to the previous section, here we relax the exponential moments assumption. Here we just require that $\mathrm{E}\, X_1^{\alpha+2} < \infty$ for some $\alpha \geq 0$. If the moment generating function of the X_is exists, Theorem 2 corrects Renyi's approximation (1) by providing a full asymptotic expansion in powers of p with an exponential error term. In other words, Theorem 2 provides rigorous support for the parametric (in $p > 0$) approximation

$$\mathrm{P}\left(S_M > \frac{x}{p}\right) \approx \exp\left(-\frac{x}{\mathrm{E}\, X_1} + \sum_{k=1}^{\infty} p^k (\xi_k - \gamma_{k+1} x)\right),$$

which is valid up to an error exponentially small as $p \searrow 0$. It is easy to see that γ_k and ξ_k depend on the first k and $(k+1)$ order moments of X_1, respectively. This suggests that, if $\mathrm{E}\, X_1^{\alpha+2} < \infty$, the approximation

$$\mathrm{P}\left(S_M > \frac{x}{p}\right) \approx \exp\left(-\frac{x}{\mathrm{E}\, X_1} + \sum_{1 \leq k \leq \alpha+1} p^k (\xi_k - \gamma_{k+1} x)\right) \qquad (28)$$

should be more accurate than (1). Another (perhaps more natural) way of obtaining a formal expression such as (28) proceeds as follows. First define

$$\phi_\alpha(\theta) = 1 + \sum_{k \leq \alpha+2} \frac{\mathrm{E}\, X^k \theta^k}{k!}.$$

Then find the smallest positive solution $\tilde{\theta}_\alpha$ to the equation

$$\phi_\alpha(\tilde{\theta}_\alpha) = \frac{1}{q}. \qquad (29)$$

Note that $\tilde{\theta}_\alpha$ exists and is well defined if $p > 0$ is sufficiently small. Finally, set

$$c_\alpha(p) = \frac{p}{q^2 \tilde{\theta}_\alpha \phi_\alpha'(\tilde{\theta}_\alpha)}. \qquad (30)$$

Using these elements, (28) is equivalent (up to quantities of order $o(p^{\alpha+1})$) to

$$\mathrm{P}\left(S_M > \frac{x}{p}\right) \approx \exp\left(-\frac{x\tilde{\theta}_\alpha}{p}\right) c_\alpha(p). \qquad (31)$$

Providing rigorous support for approximation (31) in the presence of heavy tails presents an additional mathematical complication. Note that a crucial ingredient in the proof of Theorem 2 is the existence of a root for (19). This indicates that the strategy followed in the proof of Theorem 2 is infeasible in the heavy-tailed case. Our idea is then to proceed by removing the large 'outlier'. Define the sequence $\overline{X} = (\overline{X}_k : k \geq 1)$ as $\overline{X}_k = X_k \mathbf{1}_{[0, x/p]}(X_k)$ and consider its associated random walk $\overline{S} = (\overline{S}_n : n \geq 0)$ (i.e. $\overline{S}_n = \overline{X}_1 + \cdots + \overline{X}_n$ with $\overline{S}_0 = 0$). We first argue that this 'truncation' has little impact on the distribution of the geometric sum over spatial scales of order $1/p$.

Lemma 5. *The RV* $\overline{S}_M$ *satisfies*

$$|\,\mathrm{P}(pS_M > x) - \mathrm{P}(p\overline{S}_M > x)| \leq \frac{2\,\mathrm{P}(X_1 > x/p)}{p}.$$

Proof. Note that

$$|\,\mathrm{P}(pS_M > x) - \mathrm{P}(p\overline{S}_M > x)|$$

$$\leq p\sum_{k=1}^{\infty} q^{k-1}\,\mathrm{P}\!\left(S_k > \frac{x}{p};\ \overline{S}_k \leq \frac{x}{p}\right) + p\sum_{k=0}^{\infty} q^{k-1}\,\mathrm{P}\!\left(\overline{S}_k > \frac{x}{p};\ S_k \leq \frac{x}{p}\right)$$

$$\leq 2p\sum_{k=1}^{\infty} q^{k-1}\,\mathrm{P}\!\left(\max_{n\leq k} X_n > \frac{x}{p}\right)$$

$$\leq 2p\sum_{k=1}^{\infty} kq^{k-1}\,\mathrm{P}\!\left(X_1 > \frac{x}{p}\right)$$

$$= \frac{2\,\mathrm{P}(X_1 > x/p)}{p}.$$

As an immediate corollary, we have the following result.

Corollary 1. *Suppose that* $\mathrm{E}\,X_1^{\beta} < \infty$ *for* $\beta \geq 1$, *then*

$$|\,\mathrm{P}(pS_M > x) - \mathrm{P}(p\overline{S}_M > x)| = o\!\left(\frac{p^{\beta-1}}{x^{\beta}}\right).$$

The strategy to follow in order to provide rigorous support for the validity of (31) is perhaps clear now. Specifically, taking advantage of the fact that the increments $\overline{X}_k$ have exponential moments, we want to deal with $\overline{S}_M$ following the spirit of Section 3 by making use of the uniform renewal theory. One of the basic steps that is involved in applying the same techniques used in Section 3 is solving the equation

$$\mathrm{E}\exp(\overline{\theta}\,\overline{X}) = \frac{1}{1-p}. \tag{32}$$

It is not hard to see, using the implicit function theorem, that $\overline{\theta} \sim p/\mathrm{E}\,X$ as $p \searrow 0$. In fact, we have the following proposition.

Proposition 3. *Suppose that* $\mathrm{E}\,X_1^{\alpha+2} < \infty$. *Then, for* $A > 0$, *we have*

$$\phi_{\alpha}(\theta) = \mathrm{E}\exp(\theta\overline{X}) + o(p^{\alpha+2})$$

as $p \searrow 0$ *uniformly for* $|\theta| \leq pA$. *Therefore, if* $\overline{\theta}$ *solves (32),*

$$\overline{\theta} = \theta_{\alpha} + o(p^{\alpha+2}).$$

Proof. Suppose that $|\theta| \leq pA$ for some fixed $A > 0$. Then,

$$\mathrm{E}\exp(\theta\overline{X}) = \mathrm{E}\!\left(e^{\theta X};\ X \leq \frac{1}{p}\right) + o(p^{\alpha+2}).$$

Also, note that

$$\left| E\left(e^{\theta X}; \ X \le \frac{1}{p} \right) - 1 - \sum_{1 \le k \le \alpha+2} \frac{E \, X^k \theta^k}{k!} \right|$$

$$\le o(p^{\alpha+2}) + A e^A \, p^{\lfloor \alpha+2 \rfloor + 1} \, E\left(X^{\lfloor \alpha+2 \rfloor + 1}; \ X \le \frac{1}{p} \right).$$

Now, we claim that

$$p^{\lfloor \alpha+2 \rfloor + 1} \, E\left(X^{\lfloor \alpha+2 \rfloor + 1}; \ X \le \frac{1}{p} \right) = o(p^{\alpha+2}).$$

In order to see this, we first use integration by parts to obtain

$$p \, E\left(X^{\lfloor \alpha+2 \rfloor + 1}; \ X \le \frac{1}{p} \right) = \int_0^{1/p} p s^{\lfloor \alpha+2 \rfloor + 1} \, P(X \in ds)$$

$$= - P\left(X > \frac{x}{p} \right) \left(\frac{1}{p} \right)^{\lfloor \alpha+2 \rfloor} x \tag{33}$$

$$+ (\lfloor \alpha + 2 \rfloor + 1) \int_0^1 P\left(X > \frac{s}{p} \right) \left(\frac{s}{p} \right)^{\lfloor \alpha+2 \rfloor} ds. \tag{34}$$

Expression (33) multiplied by $p^{\lfloor \alpha+2 \rfloor}$ is of order $o(p^{\alpha+2})$; hence, we just have to show that the integral in (34) multiplied by $p^{\lfloor \alpha+2 \rfloor}$ is of order $o(p^{\alpha+2})$. Equivalently, we must verify that

$$\int_0^1 P\left(X > \frac{s}{p} \right) \left(\frac{s}{p} \right)^{\alpha+2} s^{\lfloor \alpha+2 \rfloor - (\alpha+2)} \, ds \to 0 \quad \text{as } p \searrow 0,$$

which follows easily by dominated convergence since $0 \le P(X > t) t^{\lfloor \alpha+2 \rfloor} \le c < \infty$, and

$$\int_0^1 s^{\lfloor \alpha+2 \rfloor - (\alpha+2)} \, ds < \infty.$$

The rest just follows from the implicit function theorem.

Finally, we provide the precise statement of our rigorous approximation in the context of heavy-tailed increments. (We say here that a nonnegative random variable X_1 is heavy tailed if, for every $\eta > 0$, $E \exp(\eta X_1) = \infty$.)

Theorem 3. *Assume that the distribution of X_1 is strongly nonlattice and that $E \, X_1^{2+\alpha} < \infty$ for some $\alpha \ge 0$. Then,*

$$P(p S_M > x) = \exp\left(-\frac{x \tilde{\theta}_\alpha}{p} \right) c_\alpha(p) + o(p^{\alpha+1}) \tag{35}$$

as $p \searrow 0$ uniformly over x on compact sets.

The proof of this result will be provided at the end of this section.

As in the case of exponential moments, approximation (35) can be stated in power series form as we indicate next.

Corollary 2. *In the setting of Theorem 3,*

$$\mathrm{P}\!\left(S_M > \frac{x}{p}\right) = \exp\!\left(-\frac{x}{\mathrm{E}\,X_1} + \sum_{1 \le k \le \alpha+1} p^k(\xi_k - \gamma_{k+1}x)\right) + o(p^{\alpha+1}),$$

where the γ_ks and ξ_ks are defined recursively via Proposition 2.

Proof. Follows from the implicit function theorem applied to (29), and then expanding (30) in powers of p.

We are now ready to provide the proof of Theorem 3. We do so by analyzing $\mathrm{P}(p\overline{S}_M > x)$ just as we did in Theorem 2. Theorem 1 can also be applied here to obtain a suitable approximation for $\mathrm{P}(p\overline{S}_M > x)$, as our next result shows and this is the strategy that we follow next.

Proof of Theorem 3. To compute $\mathrm{P}(pS_M > x)$ we 'truncate' at level x/p, thereby introducing the $\overline{X}_k$s, where $\overline{X}_k = X_k \mathbf{1}_{[0,x/p]}(X_k)$. Then, $\overline{a}(\cdot) = \mathrm{P}(p\overline{S}_m > \cdot)$, when evaluated at x/p, satisfies

$$\exp\!\left(\overline{\theta}\frac{x}{p}\right)\overline{a}\!\left(\frac{x}{p}\right) = \int_{[0,x/p)} \exp\!\left(\overline{\theta}\!\left(\frac{x}{p} - s\right)\right) \mathrm{P}\!\left(X_1 > \frac{x}{p} - s;\ X_1 \le \frac{x}{p}\right)\overline{U}_{\overline{\theta}}(ds),$$

where $\overline{U}_{\overline{\theta}}(s) = \sum_{n=0}^{\infty}\mathrm{P}_{\overline{\theta}}(S_n \le s)$, $\mathrm{P}_{\overline{\theta}}(\cdot)$ is defined via

$$\mathrm{P}_{\overline{\theta}}(B) = q^n\,\mathrm{E}(\exp(\overline{\theta}\overline{S}_n); \mathbf{1}_{(B)})$$

for every B in the sigma-field $\sigma(\overline{X}_1, \ldots, \overline{X}_n)$, and $\overline{\theta}$ is the solution to the equation

$$\overline{\phi}(\overline{\theta}) := \mathrm{E}\exp(\overline{\theta}\overline{X}_1) = q^{-1}.$$

Next, we will show that

$$\overline{V}(s) := \overline{U}_{\overline{\theta}}(s) - \frac{s}{\mathrm{E}_{\overline{\theta}}\overline{X}_1} - \frac{\mathrm{E}_{\overline{\theta}}\overline{X}_1^2}{2\,\mathrm{E}_{\overline{\theta}}^2\overline{X}_1} - \frac{\int_s^{\infty}\int_t^{\infty}\mathrm{P}_{\overline{\theta}}(\overline{X}_1 > u)\,du\,dt}{\mathrm{E}_{\overline{\theta}}^2\overline{X}_1}, \tag{36}$$

satisfies $|\overline{V}(s)| = o(s^{-(\alpha+1)})$ as $s \nearrow \infty$ uniformly in $p > 0$ small enough. This follows from Theorem 1, as we now illustrate. (Note that the term $\overline{V}$ in (36) includes the last two terms on the right-hand side of (5).) Observe that $\overline{g}_p(\lambda) := \mathrm{E}_{\overline{\theta}}\exp(\mathrm{i}\lambda\overline{X}_1) = q\,\mathrm{E}\exp((\mathrm{i}\lambda + \overline{\theta})\overline{X}_1)$ satisfies

$$
\begin{aligned}
|\overline{g}_p(\lambda) - \mathrm{E}\exp(\mathrm{i}\lambda X_1)| &\le |\overline{g}_p(\lambda) - \mathrm{E}\exp(\mathrm{i}\lambda\overline{X}_1)| + o(p^{\alpha+2}) \\
&\le p|\mathrm{E}\exp(\mathrm{i}\lambda\overline{X}_1)| + \overline{\theta}\,\mathrm{E}\,\overline{X}_1 + o(p^{\alpha+2}) \\
&= O(p).
\end{aligned}
$$

Since X_1 is strongly nonlattice, this implies that $\overline{g}_p(\cdot)$ satisfies the uniform strongly nonlattice condition, (14). Conversely, since $\overline{\theta} = O(p)$, we find that, for all $p > 0$ small enough,

$$\mathrm{E}_{\overline{\theta}}\overline{X}_1^{\alpha+2} = q\,\mathrm{E}\exp(\overline{\theta}\overline{X}_1)\overline{X}_1^{\alpha+2} \le M\,\mathrm{E}\,\overline{X}_1^{\alpha+2} < M\,\mathrm{E}\,X_1^{\alpha+2} < \infty.$$

Theorem 1 now justifies the validity of (36). Furthermore, (36) implies that

$$\overline{a}\left(\frac{x}{p}\right) = \int_{[0,x/p)} \frac{\exp(\overline{\theta}(x/p - s))\, P(X_1 > x/p - s;\ X_1 \le x/p)}{E_{\overline{\theta}} \overline{X}_1}\, ds \tag{37}$$

$$+ \int_{[0,x/p)} \frac{\exp(\overline{\theta}(x/p - s))\, P(X_1 > x/p - s;\ X_1 \le x/p)}{E_{\overline{\theta}}^2 \overline{X}_1} \int_s^\infty P_{\overline{\theta}}(\overline{X}_1 > u)\, du\, ds \tag{38}$$

$$+ \int_{[0,x/p)} \exp\left(\overline{\theta}\left(\frac{x}{p} - s\right)\right) P\left(X_1 > \frac{x}{p} - s;\ X_1 \le \frac{x}{p}\right) \overline{V}(ds). \tag{39}$$

Let us denote by I_1, I_2, and I_3 the expressions in (37), (38), and (39), respectively. We first show that $I_3 = o(p^{\alpha+1})$. To see this, we use integration by parts, the triangle inequality, and the fact that $\overline{\theta} = O(p)$ to obtain

$$|I_3| \le \left|\overline{V}\left(\frac{x}{p}\right)\right| + M_1 \left|\int_{[0,x/p)} \overline{V}(s)\, d\exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - s;\ X_1 \le \frac{x}{p}\right)\right|$$

$$\le \left|\overline{V}\left(\frac{x}{p}\right)\right| + M_1 \left|\int_{[0,x/p)} \overline{V}(s)\, \exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - ds;\ X_1 \le \frac{x}{p}\right)\right|$$

$$+ M_1 \left|\int_{[0,x/p)} \overline{V}(s)\, \exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - s;\ X_1 \le \frac{x}{p}\right) ds\right|. \tag{40}$$

We showed earlier that $|\overline{V}(x/p)| = o(p^{\alpha+1})$. Now, observe that

$$\left|\int_{[0,x/p)} \overline{V}(s)\, \exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - ds;\ X_1 \le \frac{x}{p}\right)\right|$$

$$\le \left|\int_{[0,x/2p)} \overline{V}(s)\, \exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - ds;\ X_1 \le \frac{x}{p}\right)\right|$$

$$+ \left|\int_{[x/2p,x/p)} \overline{V}(s)\, \exp(-\overline{\theta}s)\, P\left(X_1 > \frac{x}{p} - ds;\ X_1 \le \frac{x}{p}\right)\right|$$

$$\le K_2\, P\left(X_1 > \frac{x}{2p}\right) + K_1 \max_{1/2 \le u \le 1} \left|\overline{V}\left(\frac{ux}{p}\right)\right|$$

$$= o(p^{\alpha+2}) + o(p^{\alpha+1})$$

$$= o(p^{\alpha+1}) \tag{41}$$

for some constants K_1 and K_2. The integral in (40) follows the same lines as (41). For I_2, we have

$$I_2 = \frac{1}{E_{\overline{\theta}}^2 \overline{X}_1} \int_{[0,x/p)} \exp\left(\overline{\theta}\left(\frac{x}{p} - s\right)\right) P\left(X_1 > \frac{x}{p} - s\right) \int_s^\infty P_{\overline{\theta}}(\overline{X}_1 > u)\, du\, ds + o(p^{\alpha+1}).$$

Note that

$$\frac{1}{E_{\overline{\theta}}^2 \overline{X}_1} \int_{[0,x/p)} \exp\left(\overline{\theta}\left(\frac{x}{p} - s\right)\right) P\left(X_1 > \frac{x}{p} - s\right) \int_s^\infty P_{\overline{\theta}}(\overline{X}_1 > u)\, du\, ds$$

$$\sim \frac{1}{E^2 X} \int_0^{x/p} P\left(X_1 > \frac{x}{p} - s\right) \int_s^{x/p} P(X_1 > u)\, du\, ds. \tag{42}$$

Let J be the integral in (42). We must show that $J = o(p^{\alpha+1})$. To see this just note that

$$J = \frac{1}{\mathrm{E}^2\, X} \int_0^{x/2p} \mathrm{P}\left(X_1 > \frac{x}{p} - s\right) \int_s^{x/p} \mathrm{P}(X_1 > u)\, \mathrm{d}u\, \mathrm{d}s$$

$$+ \frac{1}{\mathrm{E}^2\, X} \int_{x/2p}^{x/p} \mathrm{P}\left(X_1 > \frac{x}{p} - s\right) \int_s^{x/p} \mathrm{P}(X_1 > u)\, \mathrm{d}u\, \mathrm{d}s$$

$$\leq \frac{x\, \mathrm{P}(X_1 > x/2p)}{2p\, \mathrm{E}\, X} + \int_{x/2p}^{x/p} \mathrm{P}(X_1 > u)\, \mathrm{d}u\, \frac{1}{\mathrm{E}\, X}$$

$$\leq \frac{x\, \mathrm{P}(X_1 > x/2p)}{p\, \mathrm{E}\, X}$$

$$= o(p^{\alpha+1}),$$

which yields $I_2 = o(p^{\alpha+1})$. Finally, we analyze I_1:

$$I_1 + o(p^{\alpha+1}) = \frac{1}{\mathrm{E}_{\bar{\theta}}\, \bar{X}_1} \int_0^{x/p} \exp(\bar{\theta} u)\, \mathrm{P}(X_1 > u)\, \mathrm{d}u$$

$$= \frac{1}{\bar{\theta}\, \mathrm{E}_{\bar{\theta}}\, \bar{X}_1} \int_0^{x/p} \mathrm{P}(X_1 \geq u)\, \mathrm{d}\exp(\bar{\theta} u)$$

$$= o(p^{\alpha+1}) - \frac{1}{\bar{\theta}\, \mathrm{E}_{\bar{\theta}}\, \bar{X}_1} + \frac{1}{\bar{\theta}\, \mathrm{E}_{\bar{\theta}}\, \bar{X}_1} \int_0^{x/p} e^{\bar{\theta} u}\, \mathrm{P}(X_1 \in \mathrm{d}u)$$

$$= \frac{(1 - \mathrm{E}\exp(\bar{\theta}\bar{X}_1))}{\bar{\theta}\, \mathrm{E}(\exp(\bar{\theta}\bar{X}_1)\bar{X}_1)} + o(p^{\alpha+1}).$$

Lastly, the implicit function theorem yields

$$\frac{p}{q^2 \bar{\theta} \bar{\phi}'(\bar{\theta})} = \frac{(1 - \mathrm{E}\exp(\bar{\theta}\bar{X}_1))}{q\bar{\theta}\, \mathrm{E}(\exp(\bar{\theta}\bar{X}_1)\bar{X}_1)} = \exp\left(\sum_{k \leq \alpha+1} p^k \xi_k\right) + o(p^{\alpha+1})$$

and

$$\bar{\theta} = \sum_{k \leq \alpha+2} p^k \gamma_k + o(p^{\alpha+2}).$$

5. Geometric sums with increment distribution depending on p

As we indicated before, many settings of interest demand treatment of the case in which the increment distributions are actually changing with p. This is the case if we wish to develop CDAs for the time-in-system for the GI/G/1 queue and the probability of ruin for the renewal risk insurance process. Fortunately, Theorem 1 also provides a means to deal with the typical situations that arise in practice. To fix ideas, consider a family of probability measures $\mathcal{P} = \{\mathrm{P}_p,\, p \in [0, \bar{\delta}]$ for some $\bar{\delta} > 0\}$. Suppose that, under each P_p, the random variables $(X_k : k \geq 1)$ form an i.i.d. sequence. Also, assume that the distribution of X_1 is uniformly strongly nonlattice with respect to $\mathcal{P}$ (i.e. the characteristic functions $g_p(\lambda) = \mathrm{E}_p \exp(\mathrm{i}\theta X_1)$ satisfy condition (4)). In addition, suppose that one of the following conditions hold:

(A) for some $\eta > 0$, $\sup_{0 \leq p \leq \bar{\delta}} \mathrm{E}_p \exp(\eta X_1) < \infty$; or

(B) $\sup_{0 \leq p \leq \bar{\delta}} \mathrm{E}_p\, X_1^{2+\alpha} < \infty$ for some $\alpha \geq 0$.

Under this set of assumptions, we have the following analogue to Theorems 2 and 3.

Theorem 4. *Assume that the family* P_p, $p \in [0, \bar{\delta}]$, *is uniformly strongly nonlattice (see (4)). If condition (A) holds then there exist constants* K_1, $K_2 > 0$ *such that, for* $p > 0$ *small,*

$$\left| P_p(pS_M > x) - \exp\left(-\frac{\theta^* x}{p} + r_p(p) \right) \right| \leq K_1 \exp\left(-\frac{K_2 x}{p} \right), \qquad (43)$$

where $\theta^* = \theta^*(p)$ *solves* $\phi_p(\theta^*) := E_p \exp(\theta^* X_1) = 1/q$ *and*

$$\exp(r_p(p)) = \frac{p}{q^2 \theta^* \phi_p'(\theta^*)}.$$

Moreover, $\theta^*(p) = \sum_{k=1}^{\infty} p^k \gamma_k(p)$ *and* $r_p(p) = \sum_{k=1}^{\infty} p^k \xi_k(p)$ *(where the* $\gamma_k(p)s$ *and* $\xi_k(p)s$ *are defined in terms of the first k and $(k+1)$ moments of X_1 under P_p, respectively, as in Remark 4). Finally, if condition (B) holds then*

$$\left| P_p(pS_M > x) - \exp\left(-\frac{x}{E_p X_1} + \sum_{1 \leq k \leq \alpha+1} p^k (\xi_k - \gamma_{k+1} x) \right) \right| = o(p^{\alpha+1}). \qquad (44)$$

Proof. The proof parallels the arguments given in Theorems 2 and 3 using Theorem 1. The details are omitted.

Remark 4. Note that the γ_ks and the ξ_ks also depend on p. The previous result yields an asymptotic expansion, assuming that the problem at hand has enough structure (i.e. when an asymptotic expansion of the ξ_ks and γ_ks in powers of p can be obtained). The expansion for the distribution of the all time maximum of a random walk with small negative drift given in Blanchet and Glynn (2006) provides an example of the applicability of this result.

Remark 5. Just as we pointed out in Theorems 2 and 3, the estimate in (43) applies uniformly in x, whereas (44) holds as long as $x = x(p)$ stays bounded.

6. Geometric sums with real-valued increments

Some contexts demand looking at increment distributions that can take negative values. We refer the reader to the book by Kalashnikov (1997) for motivating applications. Our goal in this section is to show that completely analogous results to those presented in Section 3 can be obtained even if we relax the assumption of nonnegative increments. The strategy is a natural extension to that of Section 3. For simplicity, we shall assume that $E \exp(\eta |X_1|) < \infty$. Additional order-correction terms can be obtained, just as in Theorem 3 via a truncation argument.

First let us consider the case in which $E X_1 > 0$. Note that owing to the memoryless property of the geometric distribution, we obtain

$$G(x) := P(S_M > x) = p P(X_1 > x) + q P(\widetilde{Z} + X_1 > x),$$

where $S_M \overset{\text{D}}{=} \widetilde{Z}$ (where '$\overset{\text{D}}{=}$' denotes equality in distribution) and $\widetilde{Z}$ is independent of X_1. Therefore,

$$G(x) = p P(X_1 > x) + q \int_{-\infty}^{\infty} G(x - s) P(X_1 \in ds).$$

As in Section 3, let $\tilde{\theta} > 0$ be such that $\mathrm{E}\exp(\tilde{\theta}X_1) = 1/q$. Then,

$$\exp(\tilde{\theta}x)G(x) = p\int_{-\infty}^{\infty} \mathrm{P}(X_1 > x - s)\exp(\tilde{\theta}(x-s))U_{\tilde{\theta}}(\mathrm{d}s), \tag{45}$$

where $U_{\tilde{\theta}}(t) = \sum_{n=0}^{\infty}\mathrm{P}_{\tilde{\theta}}(S_n \leq t)$, and under $\mathrm{P}_{\tilde{\theta}}$ the increments of the random walk $S = (S_n\colon n \geq 1)$, namely the X_is, are i.i.d. with distribution

$$\mathrm{P}_{\tilde{\theta}}(X \in \mathrm{d}s) = q\exp(\tilde{\theta}s)\mathrm{P}(X \in \mathrm{d}s).$$

We then obtain the next result, which is the analog of Theorem 2.

Theorem 5. *Suppose that X_1 has strongly nonlattice distribution, $\mathrm{E}\,X_1 > 0$, and that $\phi(\eta) := \mathrm{E}\,\mathrm{e}^{\eta|X|} < \infty$ for some $\eta > 0$. If $x > 0$ then there exists $a > 0$ such that*

$$\mathrm{P}(pS_M > x) = \exp\left(-\frac{x\tilde{\theta}}{p} + r(p)\right) + O\left(\exp\left(-\frac{ax}{p}\right)\right) \quad as\ p \to 0,$$

where

$$\mathrm{e}^{r(p)} = \frac{p}{q^2\tilde{\theta}\phi'(\tilde{\theta})} := c(p). \tag{46}$$

If $x < 0$ then there exists $a > 0$ such that

$$\mathrm{P}(pS_M \leq x) = O(\mathrm{e}^{ax/p}).$$

Finally,

$$\mathrm{P}(S_M \leq 0) = \sum_{n=0}^{\infty}pq^{n-1}\,\mathrm{P}(S_n \leq 0)$$

is real analytic in $p \in [0, \delta]$ for some $\delta > 0$.

    Proof. Theorem 1 yields

$$U_{\tilde{\theta}}(s) = \mathbf{1}_{[0,\infty)}(s)\left(\frac{s}{\mathrm{E}_{\tilde{\theta}}\,X_1} + \frac{\mathrm{E}_{\tilde{\theta}}\,X_1^2}{\mathrm{E}_{\tilde{\theta}}^2\,X_1} + V_1(s)\right) + V_2(s)\,\mathbf{1}_{(0,\infty)}(s),$$

where $|V_1(s)|\,\mathbf{1}_{[0,\infty)}(s) + |V_2(s)|\,\mathbf{1}_{(0,\infty)}(s) = O(\mathrm{e}^{-r|s|})$ as $|s| \to \infty$ for some $r > 0$ (uniformly over $p \in [0, \delta]$ for some $\delta > 0$). Hence, (45) implies that

$$\exp\left(\tilde{\theta}\frac{x}{p}\right)\mathrm{P}(pS_M > x) = p\int_{-\infty}^{\infty}\mathrm{P}\left(X_1 > \frac{x}{p} - s\right)\exp\left(\tilde{\theta}\left(\frac{x}{p} - s\right)\right)U_{\tilde{\theta}}(\mathrm{d}s)$$

$$= p\int_{-\infty}^{x/p}\frac{\mathrm{P}(X_1 > u)\exp(\tilde{\theta}u)\,\mathrm{d}u}{\mathrm{E}_{\tilde{\theta}}\,X_1} \tag{47}$$

$$+ p\int_{0}^{\infty}\mathrm{P}\left(X_1 > \frac{x}{p} - s\right)\exp\left(\tilde{\theta}\left(\frac{x}{p} - s\right)\right)\mathrm{d}V_1(s) \tag{48}$$

$$+ p\int_{0}^{\infty}\mathrm{P}\left(X_1 > \frac{x}{p} + s\right)\exp\left(\tilde{\theta}\left(\frac{x}{p} + s\right)\right)V_2(-\mathrm{d}s). \tag{49}$$

It is not hard to see that integrals (48) and (49) are of order $O(e^{-rx/p})$ for some $r > 0$ uniformly in $p \in [0, \delta]$ for some $\delta > 0$. Hence, we obtain

$$\exp\left(\tilde{\theta}\frac{x}{p}\right) P(pS_M > x) = p \int_{-\infty}^{\infty} \frac{P(X_1 > u) \exp(\tilde{\theta}u)\, du}{E_{\tilde{\theta}} X_1} + O(e^{-rx/p})$$

for some $r > 0$. Conversely, integrating by parts we obtain

$$\int_{-\infty}^{\infty} P(X_1 > u)\tilde{\theta} \exp(\tilde{\theta}u)\, du = \int_{-\infty}^{\infty} \exp(\tilde{\theta}u) P(X_1 \in du) = \frac{1}{q}.$$

Therefore,

$$\exp\left(\tilde{\theta}\frac{x}{p}\right) P(pS_M > x) = \frac{p}{\tilde{\theta}q\, E_{\tilde{\theta}} X_1} + O(e^{-rx/p}).$$

Equation (46) is obtained by noting that $E_{\tilde{\theta}} X_1 = q\phi'(\tilde{\theta})$. The behavior of $P(S_M \leq x/p)$ for $x < 0$ can be obtained using a similar analysis to the previous one. The analyticity of $P(S_M \leq 0)$ follows using Chernoff's bounds by noting that we can choose $\rho \in (0, 1)$ for which

$$\sum_{n=0}^{\infty} pq^{n-1} P(S_n \leq 0) \leq \sum_{n=0}^{\infty} p(1 + p)^{n-1} P(S_n \leq 0)$$

$$\leq \sum_{n=0}^{\infty} p(1 + p)^{n-1} \rho^n$$

$$< \infty,$$

as long as $p \in [0, \delta]$ for $\delta > 0$ sufficiently small. Hence, the series representation for $P(S_M \leq 0)$ converges absolutely and uniformly for sufficiently small p. This concludes the proof.

7. Geometric sums with zero-mean increments

The case in which $E X_1 = 0$ introduces qualitative differences. As we discussed in our introduction, it is not hard to see that $p^{1/2} S_M \xrightarrow{W} \sigma 2^{-1/2} T$, where T follows the Laplace (or double exponential) distribution. The scaling of this weak convergence result suggests that the expansion in this case is given in powers of $p^{1/2}$. The next theorem considers $P(p^{1/2} S_M > x)$ for $x > 0$; the case in which $x = 0$ is investigated separately.

Theorem 6. *Suppose that X_1 has a strongly nonlattice distribution, $E X_1 = 0$, and that $\phi(\eta) := E e^{\eta|X|} < \infty$ for some $\eta > 0$. If $x > 0$ then there exists $a > 0$ such that*

$$P(p^{1/2} S_M > x) = \exp\left(-\frac{x\tilde{\theta}}{p^{1/2}} + r(p)\right) + O\left(\exp\left(-\frac{ax}{p^{1/2}}\right)\right) \quad \text{as } p \to 0,$$

where

$$e^{r(p)} = \frac{p}{q^2 \tilde{\theta} \phi'(\tilde{\theta})} := c(p).$$

Proof. We proceed as in the argument given at the beginning of this section to obtain

$$\exp(\tilde{\theta}x) P(p^{1/2} S_M > x) = p \int_{-\infty}^{\infty} P(X_1 > xp^{-1/2} - s) \exp(\tilde{\theta}(x - s)) U_{\tilde{\theta}}(ds), \qquad (50)$$

where $U_{\tilde\theta}(t) = \sum_{n=0}^{\infty} P_{\tilde\theta}(S_n \leq t)$, and under $P_{\tilde\theta}$ the increments of the random walk $S = (S_n : n \geq 1)$, namely the X_is, are i.i.d. with distribution

$$P_{\tilde\theta}(X \in ds) = q \exp(\tilde\theta s) P(X \in ds),$$

again, $\tilde\theta$ is chosen at the unique nonnegative root of the equation $E \exp(\tilde\theta X) = q^{-1}$. Note that $\mu_{\tilde\theta} := E_{\tilde\theta} X_1 > 0$, so the representation given in (50) is valid by virtue of standard renewal theory. We can proceed just as in the proof of Theorem 5; we just have to be more careful here because $E_{\tilde\theta} X_1 = O(p^{1/2})$. Note that the integral in (50) can be written as

$$\exp\left(\tilde\theta\frac{x}{p}\right) P(p^{1/2} S_M > x)$$

$$= p \int_{-\infty}^{\infty} P(X_1 > xp^{-1/2} - s) \exp\left(\tilde\theta\left(\frac{x}{p} - s\right)\right) U_{\tilde\theta}(ds)$$

$$= p \int_{-\infty}^{xp^{-1/2}} \frac{P(X_1 > u) \exp(\tilde\theta u) \, du}{E_{\tilde\theta} X_1} \tag{51}$$

$$+ p \int_0^{\infty} P(X_1 > xp^{-1/2} - s) \exp\left(\tilde\theta\left(\frac{x}{p} - s\right)\right) dV_1(s) \tag{52}$$

$$+ p \int_0^{\infty} P(X_1 > xp^{-1/2} + s) \exp\left(\tilde\theta\left(\frac{x}{p} + s\right)\right) V_2(-ds), \tag{53}$$

where, by virtue of Theorem 1, $|V_1(s)| = O(e^{-rs})$ as $s \to \infty$ for some $r > 0$ (uniformly over $p \in [0, \delta]$ for some $\delta > 0$) and

$$|V_2(s)| \leq K\mu_{\tilde\theta} \exp(-\mu_{\tilde\theta} r s) \quad \text{as } s \to -\infty$$

for $K, r > 0$ (uniformly over $p \in [0, \delta]$ for some $\delta > 0$). Let us denote by J_1, J_2, and J_3 the integrals in (51), (52), and (53), respectively. Note that $\tilde\theta = O(p^{1/2})$, which implies that the treatment of J_1 proceeds just as that of (47) in the proof of Theorem 5. Moreover, J_2 is also very similar to (48). The analysis of J_3 deserves special attention because the decay rate of $V_2(\cdot)$ degrades as $p \searrow 0$. Observe, using integration by parts, that, for small enough p,

$$J_3 = p \int_0^{\infty} P(X_1 > xp^{-1/2} + s) \exp(\tilde\theta(xp^{-1/2} + s)) V_2(-ds)$$

$$= -p P(X_1 > xp^{-1/2}) \exp(\tilde\theta xp^{-1/2}) V_2(0) \tag{54}$$

$$+ p\tilde\theta \int_0^{\infty} V_2(-s) \exp(\tilde\theta(xp^{-1/2} + s)) P(X_1 > xp^{-1/2} + s) \, ds \tag{55}$$

$$+ p\tilde\theta \int_0^{\infty} V_2(-s) \exp(\tilde\theta(xp^{-1/2} + s)) P(X_1 \in xp^{-1/2} + ds). \tag{56}$$

Clearly, (54) and (55) are of order $O(\exp(-rxp^{-1/2}))$ for some $r > 0$. Now, observe that (56) is bounded in absolute value by

$$p\tilde\theta \int_0^{\infty} \mu_{\tilde\theta} K \exp(-r\mu_{\tilde\theta} s) e^{\tilde\theta(xp^{-1/2} + s)} P(X_1 \in xp^{-1/2} + ds). \tag{57}$$

Applying integration by parts to the previous integral allows us to conclude that we can find constants $K_1, K_2, r_1, m > 0$ (independent of p) such that (57) is bounded by

$$K_1 \exp(-rxp^{-1/2}) + \mu_{\tilde{\theta}} p\tilde{\theta}^2 \int_0^\infty P(X_1 > xp^{-1/2} + s) K \exp(-r\mu_{\tilde{\theta}}s) \exp(\tilde{\theta}(xp^{-1/2} + s)) \, ds$$

$$\leq K_1 \exp(-rxp^{-1/2}) + K_2 \exp(\tilde{\theta}xp^{-1/2}) \int_0^\infty P(X_1 > xp^{-1/2} + um\tilde{\theta}^{-1}) \exp(-r_1 u) \, du$$

$$= O(\exp(-rxp^{-1/2})),$$

where the bound on the second line above was obtained using the change of variables $\tilde{\theta}s/m = u$ and noting that $\mu_{\tilde{\theta}}/\tilde{\theta} \to E X^2 > 0$. We therefore obtain, for some $r > 0$,

$$\exp\left(\frac{\tilde{\theta}x}{p^{1/2}}\right) P(S_M > x/p^{1/2}) = \frac{p}{\tilde{\theta}q \, E_{\tilde{\theta}} \, X_1} + O\left(\exp\left(\frac{-rx}{p^{1/2}}\right)\right).$$

Which yields, just as in Theorem 5, the existence of $a > 0$ such that

$$\exp\left(\frac{\tilde{\theta}x}{p^{1/2}}\right) P(S_M > x/p^{1/2}) = \frac{p}{q^2\tilde{\theta}\phi'(\tilde{\theta})} + O\left(\exp\left(\frac{-ax}{p^{1/2}}\right)\right).$$

In order to recover the standard weak convergence result for the double exponential RV discussed before, observe that

$$\tilde{\theta} \sim \frac{p^{1/2}2^{1/2}}{E^{1/2}(X^2)} \quad \text{and} \quad \phi'(\tilde{\theta}) \sim (E(X^2)2p)^{1/2}.$$

This implies, in particular, that

$$P(p^{1/2}S_M > x) \to \frac{\exp(-|x|2^{1/2}/E^{1/2}(X^2))}{2},$$

which is equivalent to the weak convergence result discussed before. Moreover, it is easy to see that in this case, $\tilde{\theta}(p)$ can be written as an absolutely convergent power series in $p^{1/2}$ for $p \in [0, \delta]$ with $\delta > 0$ small enough. To see this, let us write $\psi(\theta) = \log \phi(\theta)$, and note that $\tilde{\theta}$ satisfies

$$\psi(\tilde{\theta})^{1/2} = p^{1/2}\left(1 + \frac{p}{2} + \frac{p^2}{3} + \cdots\right)^{1/2}.$$

Note that $\psi(\tilde{\theta})^{1/2}$ is a real analytic function on $[0, \delta]$ for $\delta > 0$ small enough and differentiable at 0 from the right with derivative equal to $E X^2 > 0$ which yields (using the inverse function theorem) the required expansion for $\tilde{\theta}(p)$. A system of equations completely analogous to that of Proposition 2 can be easily obtained here to retrieve the coefficients in the asymptotic expansion of $\tilde{\theta}$; the details are omitted. Using an expansion for $c(p)$ in powers of $\tilde{\theta}$ yields the desired asymptotic expansion in powers of $p^{1/2}$.

While the expansion for $P(p^{1/2}S_M > x)$ follows by taking advantage of the uniform renewal theory developed in Theorem 1, the case $P(p^{1/2}S_M \leq 0)$ demands a completely different strategy. To make this computation rigorous we require, as an additional technical condition, the existence of a density for X_1. The general strongly nonlattice case should be handled by smoothing.

Proposition 4. *Suppose that the X_is possess a density and that $\mathrm{E}\exp(\theta|X_1|) < \infty$ for $\theta \in (-\varepsilon, \varepsilon)$ with $\varepsilon > 0$. Then,*

$$\mathrm{P}(p^{1/2}S_M \le 0) = \sum_{n=1}^{\infty} pq^{n-1}\,\mathrm{P}(p^{1/2}S_n \le 0)$$

admits an asymptotic expansion in powers of $p^{1/2}$ for $p \in [0, \delta]$ and some $\delta > 0$.

Proof. Without loss of generality we can assume that $\mathrm{E}\,X_1 = 1$. Note that, by Fourier inversion, we have (letting $\phi(\mathrm{i}\lambda) = \mathrm{E}\exp(\mathrm{i}\lambda X_1)$)

$$\sum_{n=1}^{\infty} pq^{n-1}\,\mathrm{P}(p^{1/2}S_n \le 0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{p\phi(\mathrm{i}\lambda p^{1/2})\,\mathrm{d}\lambda}{1 - q\phi(\mathrm{i}\lambda p^{1/2})}.$$

Now, observe that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{p\phi(\mathrm{i}\lambda p^{1/2})\,\mathrm{d}\lambda}{1 - q\phi(\mathrm{i}\lambda p^{1/2})} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\frac{p\phi(\mathrm{i}\lambda p^{1/2})}{1 - q\phi(\mathrm{i}\lambda p^{1/2})} - \frac{1}{1 + \lambda^2}\right]\mathrm{d}\lambda + \frac{1}{2}.$$

Also, observe that

$$\int_{-\infty}^{\infty}\left[\frac{p\phi(\mathrm{i}\lambda p^{1/2})}{1 - q\phi(\mathrm{i}\lambda p^{1/2})} - \frac{1}{1 + \lambda^2}\right]\mathrm{d}\lambda = -\int_{-\infty}^{\infty}\frac{1}{1 + \lambda^2}\left[\frac{1 - \phi(\mathrm{i}\lambda p^{1/2})}{1 - q\phi(\mathrm{i}\lambda p^{1/2})}\right]\mathrm{d}\lambda$$

$$= -\int_{-\infty}^{\infty}\frac{p^{1/2}}{p + \lambda^2}\,\mathrm{Re}\left[\frac{1 - \phi(\mathrm{i}\lambda)}{1 - q\phi(\mathrm{i}\lambda)}\right]\mathrm{d}\lambda.$$

The expansion now follows as in the proof of Proposition 2 of Blanchet and Glynn (2006). ∎

8. An Application to corrected diffusion approximations for the M/G/1 queue

In this section we will apply the results obtained in previous sections to develop CDAs for the time-in-system in the M/G/1 queueing model under first-in-first-out protocol; see Asmussen (2003, Chapter VIII). Using results by Siegmund (1979), a first-order CDA for the M/G/1 queue was developed by Asmussen (1984) under exponential moments. In the heavy-tailed setting (assuming the existence of five moments in the underlying processing time distributions) Asmussen and Binswanger (1997) adapted the work of Hogan (1986) to provide a first-order CDA for the time-in-system. We provide here, under weaker hypotheses, additional correction terms to the papers discussed above.

Recall that in the M/G/1 queue, customers arrive at a single-server queueing system according to a Poisson process with rate λ. The nth customer requires an amount V_n of service time. Assume that the sequence $V = (V_n: n \ge 0)$ is i.i.d. and independent of the arrival process. Suppose that

$$\mathrm{E}\,V^4 < \infty \quad \text{and} \quad \rho := \lambda\,\mathrm{E}\,V < 1.$$

It is well known (see Asmussen (2003, Chapter VIII)) that if W has the distribution of the steady-state waiting time in queue (exclusive of service) then

$$\mathrm{P}(W > x) = \rho\,\mathrm{P}(S_M > x),$$

where S_M is a geometric sum corresponding to increments possessing the distribution function

$$F(x) = \frac{1}{E\,V} \int_0^x P(V > s)\,ds$$

and geometric parameter $p = (1 - \rho)$. We are interested in obtaining an asymptotic expansion for the distribution of W as $\rho \nearrow 1$. This is the so-called heavy-traffic regime in which the system is close to 100% utilization (a setting that often arises in applications). Let $v_j = E\,V^j$. Then a straightforward application of Theorem 3 provides the asymptotic expansion

$$P((1-\rho)W > x) = \rho \exp(-x\gamma_0 + (\gamma_1 x + \xi_1)(1-\rho) + (\gamma_2 x + \xi_2)(1-\rho)^2) \\ + o((1-\rho)^2),$$

where

$$\gamma_0 = \frac{2v_1}{v_2},$$

$$v_2^3 \gamma_1 = 2v_1 v_2^2 - \frac{4v_1^2 v_3}{3},$$

$$9\gamma_2 v_2^5 = 2(8v_3^2 v_1^2 - 12v_3 v_2^2 v_1 + 9v_2^4 - 3v_4 v_2 v_1^2)v_1,$$

$$3v_2^2 \xi_1 = 3v_2^2 + 2v_3 v_1,$$

$$6v_2^4 \xi_2 = -4v_3^2 v_1^2 + 3v_2^4 + 2v_4 v_2 v_1^2 + 4v_3 v_2^2 v_1.$$

The γ_ks are also provided in Abate *et al.* (1995) for the more general GI/G/1 queue. Also, for the GI/G/1 queue, under a different parameterization of the traffic intensity, Blanchet and Glynn (2006) provided integral expressions (depending on the whole distribution of the interarrival and service times) for the ξ_ks. Our new contribution here is the explicit computation of the ξ_ks.

Acknowledgements

References

ABATE, J., CHOUDHURY, G. L. AND WHITT, W. (1995). Exponential approximations for tail probabilities in queues. I. Waiting times. *Operat. Res.* **43,** 885–901.
ALDOUS, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic.* Springer, New York.
ASMUSSEN, S. (1984). Approximations for the probability of ruin within finite time. *Scand. Actuarial J.* **1984,** 31–57.
ASMUSSEN, S. (2001). *Ruin Probabilities.* World Scientific, Singapore.
ASMUSSEN, S. (2003). *Applied Probability and Queues.* Springer, New York.
ASMUSSEN, S. AND BINSWANGER, K. (1997). Simulation of ruin probabilities for subexponential claims. *Astin Bull.* **27,** 297–318.
BLANCHET, J. AND GLYNN, P. (2006). Corrected diffusion approximations for the maximum of light-tailed random walk. *Ann. App. Prob.* **16,** 951–953.
BLANCHET, J. AND GLYNN, P. (2007). Cramer–Lundberg approximations in the absence of exponential moments: a heavy-traffic perspective. In preparation.
BOROVKOV, A. AND FOSS, S. (2000). Estimates for overshooting and arbitrary boundary by a random walk and their applications. *Theory Prob. Appl.* **44,** 231–253.
CARLSSON, H. (1983). Remainder term estimates of the renewal function. *Ann. Prob.* **11,** 143–157.

CHANG, J. AND PERES, Y. (1997). Ladder heights, Gaussian random walks and the Riemann zeta function. *Ann. Prob.* **25,** 787–802.

DURRETT, R. (2005). *Probability: Theory and Examples.* Duxbury, New York.

FUH, C. D. (2004). Uniform Markov renewal theory and ruin probabilities in Markov random walks. *Ann. Appl. Prob.* **14,** 1202–1241.

GANELIUS, T. H. (1971). *Tauberian Reminder Theorems* (Lecture Notes Math. **232**). Springer, Berlin.

HOGAN, M. L. (1986). Comment on: 'Corrected diffusion approximations in certain random walk problems'. *J. Appl. Prob.* **23,** 89–96.

KALASHNIKOV, V. (1997). *Geometric Sums: Bounds for Rare Events with Applications*. Kluwer, Dordrecht.

KEILSON, J. (1979). *Rarity and Exponentiality.* Springer, New York.

SIEGMUND, D. (1979). Corrected diffusion approximations in certain random walk problems. *Adv. Appl. Prob.* **11,** 701–719.

SIEGMUND, D. (1985). *Sequential Analysis*. Springer, New York.

STONE, C. (1965). On moment generating functions and renewal theory. *Ann. Math. Statist*. **36,** 1298–1301.