Large Deviations for the Empirical Mean of an M/M/1 Queue

Jose Blanchet, *Columbia University*
Peter Glynn, *Stanford University*
Sean Meyn, *University of Illinois*

March 12, 2011

Abstract

Let $(Q(k) : k \geq 0)$ be an $M/M/1$ queue with traffic intensity $\rho \in (0,1)$. Consider the quantity

$$S_n(p) = \frac{1}{n} \sum_{j=1}^{n} Q(j)^p$$

for any $p > 0$. The ergodic theorem yields that $S_n(p) \to \mu(p) := E[Q(\infty)^p]$, where $Q(\infty)$ is geometrically distributed with mean $\rho/(1-\rho)$. It is known that one can explicitly characterize $I(\varepsilon) > 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(S_n(p) < \mu(p) - \varepsilon\right) = -I(\varepsilon), \quad \varepsilon > 0.$$

In this paper, we show that the approximation of the right tail asymptotics requires a different logarithm scaling, giving

$$\lim_{n\to\infty} \frac{1}{n^{1/(1+p)}} \log P\left(S_n(p) > \mu(p) + \varepsilon\right) = -C(p)\, \varepsilon^{1/(1+p)},$$

where $C(p) > 0$ is obtained as the solution of a variational problem.

We discuss why this phenomenon — Weibullian right tail asymptotics rather than exponential asymptotics — can be expected to occur in more general queueing systems.

1 Introduction

The theory of large deviations for random walks is important both in its own right and as a starting point for establishing large deviations for more complex models (such as "small noise" diffusions and "slow Markov random walks"); see, for example, [6, 9, 14]. Furthermore, because of the close connection between queues and random walks, a good understanding of the large deviations for random walks is an essential ingredient in building a successful large deviations theory for queues.

A natural extension to the existing theory of large deviations for random walks is that of Markov random walks (i.e. sums of increments that are a function of a Markov chain). As in the conventional independent and identically distributed (i.i.d.) setting, such large deviations computations are of both intrinsic interest, and as a key ingredient in the theory of queues that are fed by exogenous Markov-dependent input sequences (e.g. Markov modulated arrival streams). Our goal in this paper is to study a qualitatively new phenomenon that can arise in the context of such Markov random walks, in particular those with unbounded increments. We shall provide a more detailed analysis of a phenomenon first considered in [7]. As in [7], we focus our attention on the random walk generated by the number in system of the $M/M/1$ queue.

Let $Q = (Q(k) : k \geq 0)$ be the embedded discrete time Markov chain corresponding to the number in system of an $M/M/1$ queue with arrival rate λ and service rate μ; we assume that $\rho = \lambda/\mu \in (0,1)$ so that the stability of the system is ensured. Define

$$S_n(p) := \sum_{j=1}^{n} Q(j)^p$$

for any $p > 0$.

Note that this class of integrated processes arise as a special case of the large deviations for Markov random walks of the form

$$S_n := \sum_{j=1}^{n} f(X_j),$$

where $f(\cdot)$ is real-valued and unbounded and $X = (X_j : j \geq 0)$ is a suitably regular Markov process (e.g. geometrically ergodic). We will now provide a heuristic explanation as to why the large deviations theory in this unbounded setting will often look fundamentally different both from the large deviations for conventional i.i.d. random walks and for Markov random walks with bounded increments. Recall that in the i.i.d. light-tailed setting, the most likely way in which a large deviation of order n will occur for a random walk involving n increments is through conspiratorial behavior that persists over essentially the entire time interval of length n, and its corresponding probability is roughly exponential in n. In fact, the Gibbs conditioning principle asserts that the random walk continues to have i.i.d. increments under the conditioning, with a marginal distribution that is a suitably exponentially twisted version of the original marginal distribution; see [4], Section 3.3. A similar picture typically asserts itself in the Markov random walk setting when $f(\cdot)$ is bounded. In particular, when S_n exhibits a large deviation from its equilibrium mean of order n, the conditional behavior of the process over essentially the entire interval of length n follows that of a new Markov process for which the transition function is an exponentially twisted version of the original underlying transition function of the process, and its corresponding probability is roughly exponential in n. This exponential twist is determined by solving a suitable eigenvalue problem; see [13] for the details in the finite state space case and, for example, [10] for a discussion of general state spaces. At an intuitive level, the fact that the probability is exponential in n follows from the observation that because $f(\cdot)$ is bounded, the probability distribution of roughly n increments (that are essentially independent due to mixing) must be modified, each increment providing a $O(1)$ contribution to the probability.

On the other hand, when $f(\cdot)$ is unbounded, a large deviation of order n from the mean can be achieved in $o(n)$ time steps. Because of the mixing that is present in Markov processes, this suggests that the probability of such a large deviation will be exponential with an exponent that is $o(n)$. Furthermore, the conditional dynamics of X that achieve the large deviation will not typically involve modifying the dynamics of the process over the entire time interval $[0, n]$, and will take advantage of the fact that the "cost" of pushing X into a region having high $f-$values can be amortized over the entire time scale over which X then relaxes back to equilibrium, during which. further contributions to the large deviation will be accumulated. Furthermore, if $f(\cdot)$ is unbounded above but bounded below, this intuition suggests that the large deviations that are smaller than normal will be qualitatively identical to those in the setting in which $f(\cdot)$ is bounded (e.g. large deviations probabilities that are exponential in n), while only the large deviations that are larger than typical will exhibit the qualitatively new features that can arise in the unbounded setting (which is strikingly different from what is manifested in the theory of large deviations for random walk in the conventional context). Thus, a number of qualitatively new features manifest themselves in the setting of unbounded $f(\cdot)$.

As indicated above, the $M/M/1$ queue and the associated Markov random walks $S(p) = (S_n(p) : n \geq 0)$ form an excellent vehicle for illustrating these qualitatively new behaviors. Since we have assumed that the $M/M/1$ queue is stable we have that $Q(n) \Longrightarrow Q(\infty)$ as $n \to \infty$, where $Q(\infty)$ is geometric with mean $\rho/(1-\rho)$, so that $\mu(p) := EQ(\infty)^p < \infty$ for each $p > 0$. Note first that $f_p(x) := |x|^p$ is bounded below, so our above discussion suggests that large deviations for $S_n(p)$ that are order n smaller than $n\mu(p)$ will follow the same pattern as in the setting of bounded $f(\cdot)$. This has previously been verified in (see [11], following [10]), in which it was established that for $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log P_0 \left(S_n(p)/n < \mu(p) - \varepsilon\right) = -I(\varepsilon), \tag{1}$$

where $I(\varepsilon) > 0$ can be explicitly computed. On the other hand, for large deviations that are larger than $n\mu(p)$, observe that the geometric stationary distribution for Q implies that the probability of exceeding n^r within a busy cycle is exponential in a term of order n^r, thereby contributing an $f-$value to the Markov random walk S_p of order n^{pr}. Once Q has achieved a level of order n^r, it takes a time of order n^r for the system to relax back to equilibrium, during which the typical value of $f_p(Q(\cdot))$ continues to be of order n^{pr}, thereby yielding a total contribution to $S(p)$ of order n^{r+pr}. Thus, in order to produce a large deviation in $S_n(p)$ of order n above $n\mu(p)$, we should choose $r = 1/(p+1)$, leading to an associated large probability that is exponential in $n^{1/(p+1)}$. Rigorously verifying this intuition is the main contribution of this paper.

Theorem 1 *For each $p > 0$, there exists $C(p) > 0$ such that the following limit holds for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \frac{1}{n^{1/(p+1)}} \log P_0 \left(S_n(p)/n > \mu(p) + \varepsilon\right) = -C(p)\, \varepsilon^{1/(p+1)}. \tag{2}$$

The function $C(p)$ is characterized in Section 2 — see (11).

The limit in (2) for the case $p = 1$ might be conjectured from the results of [7]. It follows from the main results there that there is a function $K\colon (0,\infty) \to (0,\infty]$ such that for positive δ,

$$\lim_{n\to\infty} \frac{1}{n} \log P_0 \left(S_n(1)/n > \mu\left(1\right) + n\delta \right) = -K(\delta)\,. \tag{3}$$

See also Proposition 11.3.4 of [12]. For the $M/M/1$ queue, one can show that $K(\delta) \approx c\sqrt{\delta}$ for δ sufficiently close to zero and some $c > 0$ (as illustration, see Figure 7 of [7]). If formally one sets $\delta_n = \varepsilon/n$ in (3), we would obtain, for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \log P_0 \left(S_n(1)/n > \mu\left(1\right) + \varepsilon \right) = -c\sqrt{\varepsilon} \tag{4}$$

This coincides with the form of (2) in the case $p = 1$ and $c = C\left(1\right)$.

The rigorous verification of (4) was posed as an open problem in the Workshop on Simulation of Networks organized at the Newton Institute in the Summer of 2010. The motivation from a simulation perspective relates to the speed of convergence of steady-state estimation estimators.

Note that a key property of the $M/M/1$ queue is that it relaxes to equilibrium slowly. It is this property that allows us to "amortize" the effort of pushing the process Q to level n^r over a time scale of sufficient duration so as to produce the desired large deviation. This amortization strategy would fail if the process were one that relaxes to equilibrium so rapidly that only a small contribution to the area under $f(X(\cdot))$ ensued. As a consequence, when $f(\cdot)$ is unbounded, it is still possible that the interaction of $f(\cdot)$ with the dynamics of X can be such the conventional large deviations explanation (i.e. modifying the dynamics of X over an interval of length n, thereby producing a large deviation that is exponential in n) holds. For example, this is the case for random walks with light-tailed i.i.d. increments having finite support. A more interesting such example (in the context of the natural continuous time analogue) is that of a mean-reverting Ornstein-Uhlenbeck process X with $f(x) = x$. (However, even in this setting, the conventional explanation can break down with $f(x) = x^p$ with p large enough.)

In computing large deviations for Markov dependent random walks (e.g. $S_n = f(X_0) + \ldots + f(X_n)$, where $X = (X_k : k \geq 0)$ is a Markov chain), we have a reasonably clear understanding of what this theory looks like for bounded $f(\cdot)$ and geometrically ergodic chains. In particular, the theory in this context looks similar like that in the setting of light-tailed independent and identically distributed (i.i.d.) increments. In particular, the large deviations asymptotics (in logarithmic scale), both lower and upper bounds, involve normalizing by n (with a common rate function). At an intuitive level, this makes perfect sense; to create a deviation from typical behavior of order n, the cheapest way is to modify the dynamics of each increment over a time scale of order n (since the largest possible contribution to the sum is $O(1)$ in the bounded setting). The cost of modifying the likelihood of such a path is then exponential in n. Furthermore, the conditional dynamics (conditioned on the large deviation) involve a time-homogeneous Markov chain (with a transition kernel that is suitably exponentially twisted).

However, the story looks very different in the setting of unbounded $f(\cdot)$, and the $M/M/1$ queue represents perhaps the simplest example in which these new phenomena present themselves. In this case, the contribution from a given summand may no longer be $O(1)$; one may (and actually one does, as we will see) get cheaper paths corresponding to the large deviation by forcing the chain into a set where the values of $f(\cdot)$ are large (i.e. increasing in the parameter n), so that we can then generate a large deviation of order n by modifying the dynamics over a time scale that is $o(n)$. As a consequence, one can potentially obtain a large deviation of order n with a probability that is exponential of order $o(n)$ (i.e. large deviations asymptotics, in logarithmic scale, normalized by a function that is $o(n)$). Furthermore, the conditional dynamics can now involve non-time homogeneous dynamics (over a time scale of order $o(n)$) that are essentially randomly and uniformly initiated at some time within the interval $[0, n]$. We shall flesh out the details behind this story in the context of the $M/M/1$ queue.

The rest of the paper is devoted to the proof of Theorem 1.

2 Proof of Theorem 1

Our strategy is to relate the limit in (2) to a large deviations problem involving heavy-tailed random variables. For simplicity we concentrate on the case $p = 1$. Given $Q(0) = 0$, we define $T_j = \inf\{k > T_{j-1} : Q(k) = 0\}$, with $j = 1, 2, \ldots$ and $\tau_j = T_j - T_{j-1}$ for $j \geq 1$ and $T_0 = 0$. Then $Q(\cdot)$ is regenerative with respect to the sequence $\{T_j : j \geq 0\}$, which in turn induces a renewal process defined via $N(n) = \max\{k \geq 0 : T_k \leq n\}$. We write

$$S_n(p) = \sum_{j=1}^{N(n)} \sum_{k=T_{j-1}}^{T_j - 1} Q(j)^p + \sum_{k=T_{N(n)}}^{n} Q(j)^p .$$

Now, define

$$Y_j(p) = \sum_{k=T_{j-1}}^{T_j - 1} Q(j)^p = \int_{T_{j-1}}^{T_j} Q(\lfloor s \rfloor)^p \, ds, \tag{5}$$

and note that the $Y_j(p)$'s are i.i.d. Moreover, observe that $\mu(p) = EQ(\infty)^p = EY_1(p) / E\tau_1$ and therefore

$$P_0(S_n(p) > n\mu(p) + n\varepsilon) \approx P_0\Big(\sum_{j=1}^{\lfloor n/E\tau_1 \rfloor} (Y_j(p) - \mu(p)) + \sum_{k=T_{N(n)}}^{n} Q(j)^p > n\varepsilon\Big). \tag{6}$$

In our discussion here we use "$\approx$" non-rigorously, but all the statements are proved in the sense of logarithmic asymptotics as $n \to \infty$, which is ultimately the statement given in Theorem 1.

Now, since n is large, then the distribution of $Q(n)$ is close to that of $Q(\infty)$, which we denoted by $\pi(\cdot)$. In turn, since the $M/M/1$ queue is reversible, then the current cycle in progress can be interpreted as the area enclosed starting from steady-state until hitting zero. In other words, we can write the approximation

$$P_0(S_n(p) > n\mu(p)) \approx P_\pi\Big(Y_0(p) + \sum_{j=1}^{\lfloor n/E\tau_1 \rfloor} (Y_j(p) - \mu) > n\mu + n\varepsilon\Big), \tag{7}$$

where

$$Y_0(p) = \sum_{j=0}^{T_0 - 1} Q(j)^p = \int_0^{T_0} Q(\lfloor s \rfloor)^p \, ds, \text{ and } T_0 = \inf\{k \geq 0 : Q(k) = 0\}.$$

The rigorous justification behind the approximation in (7) starting from the steady-state distribution involves the following result (see [2], we provide a proof at the end of this section in order to make our exposition self-contained).

Lemma 1 *For any selection of measurable sets $A_1, \dots, A_n$*

$$P_0 \left(Q\left(k\right) \in A_k : 1 \leq k \leq n\right)$$
$$= \frac{1}{\pi\left(0\right)} P_\pi \left(Q\left(0\right) \in A_n, Q\left(1\right) \in A_{n-1}, \dots, Q\left(n-1\right) \in A_1, Q\left(n\right) = 0\right).$$

Returning to (7). Intuitively,

$$Y_0\left(p\right) \approx \Theta(\sum_{j=1}^{Q(\infty)} j^p) = \Theta(Q\left(\infty\right)^{p+1}).$$

In turn, $Q\left(\infty\right)^{p+1}$ has Weibullian-type tails with index $1/(p+1)$, and therefore $Y_0\left(p\right)$ is expected to have Weibullian tails with index (or shape parameter) $1/(p+1)$. A similar argument follows for the random variables $Y_1\left(p\right), Y_2\left(p\right), ...$, which actually have lighter tails than $Y_0\left(p\right)$ (because $Y_j\left(p\right)$ with $j \geq 1$ constitutes an accumulated area starting from the origin to the first time the process returns to the origin, whereas $Y_0\left(p\right)$ is the area starting from steady state). We now can take advantage of the way in which large deviations occur for the sum of mean-zero heavy-tailed random variables, which states that the large deviations behavior arises due to the contribution to a single large jump; this in particular explains the scaling $n^{1/(p+1)}$ that appears in (2). In order to make the principle of the single large jump rigorous we take advantage of the following result which is proved in the appendix at the end of this section.

Proposition 1 *Consider a sequence $\{Z_1, Z_2, \dots\}$ of mean zero i.i.d. random variables and assume that Z_0 is independent of the Z_j's for $j \geq 1$. Moreover, suppose that $E(Z_1^2) + E\left(Z_0^2\right) < \infty$ and that*

$$P\left(Z_0 > t\right) = \exp\left(-ct^\beta + o\left(t^\beta\right)\right) \tag{8}$$

as $t \to \infty$ for some $c \in (0, \infty)$ and $\beta \in (0, 1)$. Finally, suppose that there exists $\kappa > 0$ such that

$$P\left(Z_j > t\right) \leq \kappa P\left(Z_0 > t\right), \tag{9}$$

then for any $a > 0$

$$\log P\left(Z_0 + Z_1 + \dots + Z_n > an\right) \sim \log P\left(Z_0 > an\right)$$

as $n \to \infty$.

Although the large deviations theory for heavy-tailed (more precisely, subexponential random variables) is well developed; see for instance, [15], [3], [5] and the textbooks of [8] and [1]. The subexponential or semiexponential property in our current setting is difficult to verify. If we could verify some of these properties we might be able to obtain more precise results than just logarithmic asymptotics. Nevertheless, the assumption on the tail of Z_0 as indicated in (8) is enough for the issue that is of main concern for us, namely, logarithmic asymptotics as indicated by (2). This is why we cannot apply the existing theory directly and we needed to develop the corresponding asymptotics in Proposition 1.

Applying then the principle of the largest jump from heavy-tailed analysis yields

$$\begin{aligned} P_0(S_n > n\mu) &\approx P_\pi(\sum_{k=0}^{T_0-1} Q\left(j\right) > n\varepsilon) \\ &= P_\pi(\frac{1}{n}\int_0^{T_0} Q\left(\lfloor s \rfloor\right) ds > \varepsilon) \\ &= P_\pi(\int_0^{T_0/n^{1/2}} \frac{Q\left(\lfloor un^{1/2} \rfloor\right)}{n^{1/2}} du > \varepsilon) \end{aligned}$$

The right hand side is easy to evaluate using Laplace's principle and standard large deviations analysis given that now the initial condition can be chosen to be $Q\left(0\right) = \lfloor ny^{1/(p+1)} \rfloor$ with probability $\pi\left(\lfloor ny^{1/(p+1)} \rfloor\right) = \rho^{\lfloor ny^{1/(p+1)} \rfloor}(1-\rho)$ because now the random variable of interest has been expressed as a functional of a properly scaled (by $n^{1/(p+1)}$ in time and space) queue length process that satisfies a large deviations principle with rate $n^{1/(p+1)}$. In particular, evaluating the constant $C\left(p\right)$ in (2) is equivalent to obtaining logarithmic asymptotics for the tail of $Y_0\left(p\right)$ assuming that $Q\left(0\right)$ follows the steady-state distribution. Now, suppose that $Q\left(0\right) = y > 0$, then the distribution of $(Q\left(k\right) : k \leq T_0)$ coincides with that of $(R(k) : k \leq \mathcal{T})$, where $R\left(\cdot\right)$ is the random walk

$$R\left(k+1\right) = R\left(k\right) + W\left(k+1\right),$$

where $P(W\left(k\right) = 1) := r(\lambda) := \lambda/(\lambda+\mu) = 1 - P\left(W\left(k\right) = -1\right)$, $R\left(0\right) = y$ and $\mathcal{T} = \inf\{k \geq 0 : R\left(k\right) = 0\}$. Note that $E[W\left(1\right)] = -(\mu-\lambda)/(\mu+\lambda) < 0$.

Recall that the local large deviations rate function of $R(\cdot)$ is given by

$$
\begin{aligned}
I(z) &= \sup_{\theta}[\theta z - \log E\, exp(\theta W(k))] \\
&= \sup_{\theta}[\theta z - \log(exp(\theta) p(\lambda) + exp(-\theta)\, q(\lambda)].
\end{aligned}
$$

Mogulskii's theorem (see [4]) establishes that the sequence $\{n^{-1}R(\lfloor n\cdot\rfloor)\}$ satisfies a large deviations principle in the space of $D[0,\infty)$ with the standard Skorokhod topology with a rate function equal to

$$
J(x(\cdot)) = \int_0^\infty I(\dot{x}(s))\, ds \tag{10}
$$

where $\dot{x}(s)$ denotes the derivative of the function $x(\cdot)$ (the rate function $J(x(\cdot))$ evaluated at a function $x(\cdot)$ that is *not* absolutely continuous is set equal to infinity). Our characterization of the tail of $Y_0(p)$ given that $Q(0)$ follows the steady-state distribution takes advantage of the associated sample path large deviations for the random walk. This is precisely the content of the next result, which is proved in the appendix.

Proposition 2 *Let $C(p)$ be the solution of the variational problem of minimizing*

$$
y \log(1/\rho) + \int_0^\upsilon I\left(\dot{\phi}(s)\right) ds \tag{11}
$$

subject to $\phi(0) = y$, $\phi(s) > 0$ for all $s \in (0,\upsilon)$ and $\phi(\upsilon) = 0$; the minimization is performed over y and over functions $\phi(\cdot)$ that are absolutely continuous and that satisfy the constraints indicated earlier and also that $\int_0^\upsilon \phi(s)^p\, ds \geq 1$. Then, we have that

$$
\lim_{t\to\infty} \frac{1}{t^{1/(p+1)}} \log P_\pi(Y_0(p) > t) = -C(p).
$$

We now are ready to provide the rigorous details behind the previous outline. We prove the lower and upper bounds leading to the statement of Theorem 1 separately.

2.1 Proof of Theorem 1: The Lower Bound

First define $\overline{Q}(j) = Q(j) - \mu(p)$ and put

$$\overline{S}_n(p) = S_n(p) - n\mu(p). \tag{12}$$

We have, owing to Lemma 1,

$$\begin{aligned} &P_0\left(S_n(p) > n(\mu(p) + \varepsilon)\right) \\ &= P_0\left(\overline{S}_n(p) > n\varepsilon\right) = \frac{1}{\pi(0)} P_\pi\left(\overline{S}_n(p) > n\varepsilon, Q(n) = 0\right). \end{aligned} \tag{13}$$

Note that the initial position of the process given in the last equality of the previous display is not the origin any longer. So, the induced regenerative process is now delayed. We set $T_0 = \inf\{k \geq 0 : Q(k) = 0\}$ (note that if $Q(0) = 0$, $T_0 = 0$, we are consistent with our notation introduced for the non-delayed process given in the Introduction).As before, we put $T_j = \inf\{k > T_{j-1} : Q(k) = 0\}$ and

$$N(n) = \max\{k : T_k \leq n\}.$$

Observe that (13) induces the representation

$$\overline{S}_n(p) = \sum_{k=0}^{T_0 - 1} \overline{Q}(k) + \sum_{j=1}^{N(n)} \sum_{i=T_{j-1}}^{T_j - 1} \overline{Q}(k).$$

We let

$$Z_0 = \sum_{k=0}^{T_0 - 1} \overline{Q}(k), \text{ and } Z_j = \sum_{i=T_{j-1}}^{T_j - 1} \overline{Q}(k), \ j = 1, 2, \ldots,$$

so that the equality in (13) becomes equivalent to

$$P_0\left(S_n(p) > n\mu(p) + n\varepsilon\right) = \frac{1}{\pi(0)} P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon, Q(n) = 0). \tag{14}$$

Now, let $\varepsilon_0 > 0$ be arbitrary and note that

$$
\begin{aligned}
&P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon, Q(n) = 0) \\
&\geq P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon, |\sum_{j=1}^{N(n)} Z_j| \leq \varepsilon_0 n, Q(n) = 0) \\
&\geq P_\pi(Z_0 > n(\varepsilon + \varepsilon_0), |\sum_{j=1}^{N(n)} Z_j| \leq n\varepsilon_0, Q(n) = 0)
\end{aligned}
$$

To see that we actually have

$$
\begin{aligned}
\lim_{n\to\infty} \frac{1}{n^{p/(1+p)}} \log P_\pi \Big\{ Z_0 \geq n\varepsilon,\ \Big|\sum_{j=1}^{N(n)} Z_j\Big| \leq n\varepsilon_0,\ Q(n) = 0 \Big\} \\
= -C(p)(\varepsilon + \varepsilon_0)^{p/(1+p)},
\end{aligned}
\tag{15}
$$

we take advantage of the large deviations analysis behind the proof of Proposition 2. Note that

$$
\begin{aligned}
P_\pi(Z_0 > n(\varepsilon + \varepsilon_0)) &= P_\pi(Y_0(p) - T_0\mu(p) > n(\varepsilon + \varepsilon_0)) \\
&\leq P_\pi(Y_0(p) > n(\varepsilon + \varepsilon_0))
\end{aligned}
$$

and also note that for each $\varepsilon_1 > 0$ we have that

$$
\begin{aligned}
P_\pi(&Y_0(p) - T_0\mu(p) > n(\varepsilon + \varepsilon_0)) \\
&\geq P_\pi(Y_0(p) > n(\varepsilon + \varepsilon_0 + \varepsilon_1), T_0\mu(p) \leq n\varepsilon_1) \\
&= P_\pi(Y_0(p) > n(\varepsilon + \varepsilon_0 + \varepsilon_1)) - P_\pi(T_0\mu(p) > n\varepsilon_1).
\end{aligned}
$$

It is easy to see that T_0 has a finite moment generating function and therefore

$$
P_\pi(T_0\mu(p) > n\varepsilon_1) = O(\exp(-cn\varepsilon_1))
$$

for some $c > 0$. Since $\varepsilon_1 > 0$ is arbitrary, by Proposition 2 we conclude that

$$
\lim_{n\to\infty} \frac{\log P_\pi(Z_0 > n(\varepsilon + \varepsilon_0))}{\log P_\pi(Y_0(p) > n(\varepsilon + \varepsilon_0))} = 1. \tag{16}
$$

Now, the large deviations analysis behind Proposition 2, in particular the associated calculus of variations problem, indicates that conditional on $Z_0 > n(\varepsilon + \varepsilon_0)$, $T_0 = O\left(n^{1/(p+1)}\right)$ with very high probability. It then follows that

$$\lim_{n\to\infty} P_\pi(|\sum_{j=1}^{N(n)} Z_j| \leq n\varepsilon_0, Q(n) = 0 | Z_0 > n(\varepsilon + \varepsilon_0)) = \pi(0),$$

which, together with (16), yields (15).

2.2 Proof of Theorem 1: The Upper Bound

Our departing point is equation (14), which combined with (13) implies that

$$P_0\left(\overline{S}_n(p) > n\varepsilon\right) \leq \frac{1}{\pi(0)} P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon).$$

In turn, we have that for some $c > 0$, and each $\varepsilon_0 > 0$,

$$\begin{aligned} & P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon) \\ & \quad = P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon, |N(n) - n/E\tau_1| \leq n\varepsilon_0) + O\left(e^{-nc\varepsilon_0}\right) \end{aligned} \tag{17}$$

Moreover, we have that

$$\begin{aligned} & P_\pi(Z_0 + \sum_{j=1}^{N(n)} Z_j > n\varepsilon, |N(n) - n/E\tau_1| \leq n\varepsilon_0) \\ & \quad \leq \sum_{k=-n\varepsilon_0}^{n\varepsilon_0} P_\pi(Z_0 + \sum_{j=1}^{n/E\tau_1 + k} Z_j > (n/E\tau_1)\,\varepsilon E\tau_1). \end{aligned}$$

Using Proposition 1 we conclude that for each $k \leq \varepsilon_0 n$

$$\begin{aligned} & \limsup_{n\to\infty} \frac{1}{n^{p/(1+p)}} \log P_\pi(Z_0 + \sum_{j=1}^{n/E\tau_1 + k} Z_j \geq (n/E\tau_1)\,\varepsilon E\tau_1) \\ & \quad \leq \limsup_{n\to\infty} \frac{1}{n^{p/(1+p)}} \log P_\pi(Z_0 \geq n(1/E\tau_1 - \varepsilon_0)\,\varepsilon E\tau_1) \\ & \quad \leq -C(p)\,(\varepsilon - \varepsilon_0 \varepsilon E\tau_1)^{p/(1+p)}. \end{aligned}$$

Therefore, we conclude from these observations that

$$\limsup_{n\to\infty} \frac{1}{n^{1/(p+1)}} \log P_0 \left(\overline{S}_n(p) > n\varepsilon\right) \leq -C\,(p)\,(\varepsilon - \varepsilon_0 \varepsilon E\tau_1)^{-1/(p+1)}.$$

Since $\varepsilon_0 > 0$ is arbitrary, we arrive at the result.

2.3 Appendix: Proof of Auxiliary Results

Proof of Lemma 1. Let $K\,(i,j)$ be the Markov transition matrix of the process Q; in particular,

$$K\,(i,j) = \begin{cases} \lambda/(\lambda+\mu) & i>0, j=i+1 \\ \mu/\,(\lambda+\mu) & i>0, j=i-1 \\ 1 & i=0 \end{cases},$$

and $K\,(i,j) = 0$ otherwise. It is well known that Q is reversible, so

$$\pi\,(i)\,K\,(i,j) = \pi\,(j)\,K\,(j,i)$$

and then

$$\begin{aligned} &\pi\,(0)\,P_0\,(Q\,(k) \in A_k : 1 \leq k \leq n) \\ &= \sum_{q_1,q_2,\ldots,q_n} \pi\,(0)\,K\,(0,q_1)\ldots K\,(q_{n-1},q_n)\,I(q_k \in A_k : 1 \leq k \leq n) \\ &= \sum_{q_1,q_2,\ldots,q_n} K\,(q_1,0)\,\pi\,(q_1)\ldots K\,(q_{n-1},q_n)\,I(q_k \in A_k : 1 \leq k \leq n) \\ &\ldots \\ &= \sum_{q_1,q_2,\ldots,q_n} \pi\,(q_n)\,K\,(q_n,q_{n-1})\ldots K\,(q_2,q_1)\,K\,(q_1,0)\,I(q_k \in A_k : 1 \leq k \leq n). \end{aligned}$$

If we write $i_k = q_{n-k}$ for $k = 0,1,\ldots,n$ we obtain

$$\begin{aligned} &\pi\,(0)\,P_0\,(Q\,(k) \in A_k : 1 \leq k \leq n) \\ &= \sum_{i_0,i_1,\ldots,i_n=0} \pi\,(i_0)\,K\,(i_0,i_1)\ldots K\,(i_{n-1},i_n)\,I(i_{n-k} \in A_k : 1 \leq k \leq n). \end{aligned}$$

The right hand side is just

$$P_\pi\,(Q\,(0) \in A_n, Q\,(1) \in A_{n-1}, \ldots, Q\,(n-1) \in A_1, Q\,(n) = 0)\,.$$

Dividing by $\pi(0)$ we complete the proof of the proposition. ■

Proof of Proposition 1. We let us introduce $\varepsilon_0 > 0$ to be selected later and let us write $\delta_n = \varepsilon_0/\log(n)$. Note that

$$\begin{aligned}&P\left(Z_0 + Z_1 + \cdots + Z_n > an\right)\\ &= P(Z_0 + Z_1 + \cdots + Z_n > an, \cup_{j=0}^{n}\{Z_j > an^{1-\delta_n}\})\\ &+P(Z_0 + Z_1 + \cdots + Z_n > an, \cap_{j=0}^{n}\{Z_j \leq an^{1-\delta_n}\})\,. \end{aligned} \tag{18}$$

Now, we have that for any $\varepsilon_1 > 0$,

$$\begin{aligned}&P(Z_0 > (a+\varepsilon_1)n)P(|Z_1 + \cdots + Z_n| \leq n\varepsilon_1)\\ &= P(Z_0 > (a+\varepsilon_1)n, |Z_1 + \cdots + Z_n| \leq n\varepsilon_1)\\ &\leq\ P(Z_0 + Z_1 + \cdots + Z_n > an, \cup_{j=0}^{n}\{Z_j > an^{1-\delta_n}\})\\ &\leq P(\cup_{j=0}^{n}\{Z_j > an^{1-\delta_n}\}) \leq (n+1)\kappa P(Z_0 > an^{1-\delta_n}).\end{aligned}$$

Therefore, since $EZ_1 = 0$ we conclude that

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n^{\beta}}\log P\big(Z_0 + Z_1 + \cdots + Z_n > an, \bigcup_{j=0}^{n}\{Z_j > an^{1-\delta_n}\}\big)&\\ \leq -ca^{\beta}\exp\left(-\beta\varepsilon_0\right)&\end{aligned} \tag{19}$$

and the corresponding lower bound,

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n^{\beta}}\log P\big(Z_0 + Z_1 + \cdots + Z_n > an, \cup_{j=0}^{n}\big\{Z_j > an^{1-\delta_n}\big\}\big)&\\ \geq -c\left(a+\varepsilon_1\right)^{\beta}.&\end{aligned} \tag{20}$$

Later on we will argue that $\varepsilon_0, \varepsilon_1 > 0$ can be chosen arbitrarily small, but before we do that first let us provide an upper bound for the second term in the right hand side of (18).

Without loss of generality we can assume that $EZ_0 = 0$, if not, we redefine $Z_0 - EZ_0$ and analyze

$$P(Z_0 - EZ_0 + Z_1 + \cdots + Z_n > an - EZ_0, \cap_{j=0}^{n}\{Z_j \leq an^{1-\delta_n} - EZ_0\})\,,$$

the whole analysis to be presented would be identical, just keeping track of EZ_0 which asymptotically does not play any role; note that given the form of (8), assumption (23) would still hold after possibly redefining the constant K.

Now, define $\theta_n = \Delta n^{\gamma_n}$ such that $\gamma_n < 1-\delta_n$ and

$$\begin{aligned}
&P(Z_0 + Z_1 + \cdots + Z_n > an, \cap_{j=0}^{n}\{Z_j \le an^{1-\delta_n}\}) \\
&= P((Z_0 + Z_1 + \cdots + Z_n)/n^{1-\delta_n} > an^{\delta_n}, \cap_{j=0}^{n}\{Z_j \le an^{1-\delta_n}\}) \\
&\le (E\exp\left(\theta_n Z_0/n^{1-\delta_n}\right) I(Z_0 \le an^{1-\delta_n})) \\
&\times \left(E\exp\left(\theta_n Z_1/n^{1-\delta_n}\right) I(Z_1 \le an^{1-\delta_n})\right)^n \exp\left(-\theta_n a n^{\delta_n}\right).
\end{aligned}$$

Our strategy is to show that if Δ and γ_n are selected appropriately then the contribution of the term

$$\prod_{j=0}^{n}(E\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) I(Z_j \le an^{1-\delta_n})) = O(1) \tag{21}$$

and that we can find a constant $\kappa' > 0$ such that

$$\exp\left(-\theta_n a n^{\delta_n}\right) \le \kappa' \exp\left(-a^{\beta} c n^{\beta}(1-\varepsilon_3)\right)$$

for an arbitrary $\varepsilon_3 > 0$. To execute this strategy we first provide a Taylor development for the expectations in the product (21). Note that

$$\begin{aligned}
&E\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) I(Z_j \le an^{1-\delta_n}) \\
&= P(Z_j \le an^{1-\delta_n}) + \frac{\theta_n}{n^{1-\delta_n}} E Z_j I(Z_j \le an^{1-\delta_n}) \\
&+ \frac{\theta_n^2}{2n^{2(1-\delta_n)}} E\{\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) Z_j^2 I(Z_j \le an^{1-\delta_n})\}.
\end{aligned}$$

Note that

$$\begin{aligned}
&E\{\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) Z_j^2 I(Z_j \le an^{1-\delta_n})\} \\
&\le E\{\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) Z_j^2 I(0 \le Z_j \le an^{1-\delta_n})\} \\
&+ E\{Z_j^2 I(0 \le Z_j < 0)\}.
\end{aligned}$$

Now we have that

$$
\begin{aligned}
&E\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) Z_j^2 I(0\le Z_j\le an^{1-\delta_n})\\
&= E\int_{-Z_j}^{\infty}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(-\theta_n u/n^{1-\delta_n}\right) Z_j^2 I(0\le Z_j\le an^{1-\delta_n})du\\
&= E\int_{-\infty}^{\infty}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(-\theta_n u/n^{1-\delta_n}\right) Z_j^2 I(0\le Z_j\le an^{1-\delta_n}, -u\le Z_j)du\\
&= \int_{-\infty}^{\infty}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(-\theta_n u/n^{1-\delta_n}\right) EZ_j^2 I(-u\vee 0\le Z_j\le an^{1-\delta_n})du\\
&= EZ_j^2 I(0\le Z_j\le an^{1-\delta_n})\\
&+ \int_{-an^{1-\delta_n}}^{0}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(-\theta_n u/n^{1-\delta_n}\right) EZ_j^2 I(-u\le Z_j\le an^{1-\delta_n})du\ .
\end{aligned}
$$

Observe that

$$
\begin{aligned}
&\int_{-an^{1-\delta_n}}^{0}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(-\theta_n u/n^{1-\delta_n}\right) EZ_j^2 I(-u\le Z_j\le an^{1-\delta_n})du\\
&= \int_{0}^{an^{1-\delta_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(t\le Z_j\le an^{1-\delta_n})dt\\
&= \int_{0}^{an^{1-\delta_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j>t)dt\\
&- EZ_j^2 I(Z_j>an^{1-\delta_n})(\exp(\theta_n a)-1)\\
&\le \int_{0}^{an^{1-\delta_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j>t)dt\ .
\end{aligned}
$$

Now, let us write

$$
\begin{aligned}
&\int_{0}^{an^{1-\delta_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j>t)dt\\
&= \int_{0}^{an^{1-\delta_n-\gamma_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j>t)dt\\
&+ \int_{an^{1-\delta_n-\gamma_n}}^{an^{1-\delta_n}}\frac{\theta_n}{n^{1-\delta_n}}\exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j>t)dt\ .
\end{aligned}
$$

We analyze the previous two integrals, first we note that

$$\int_0^{an^{1-\delta_n-\gamma_n}} \frac{\theta_n}{n^{1-\delta_n}} \, exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j > t)\, dt$$
$$\leq \frac{\theta_n}{n^{1-\delta_n}} \, exp\,(a\Delta) \int_0^{\infty} EZ_j^2 I(Z_j > t)\, dt.$$

For the second integral, we use assumption (9), which implies that for any $\varepsilon_2 > 0$ there exists $K(\varepsilon_2) > 0$ such that for $t \geq 0$

$$EZ_j^2 I(Z_j > t) \leq K(\varepsilon_2) \exp\left(-c(1-\varepsilon_2)t^{\beta}\right).$$

Therefore,

$$\int_{an^{1-\delta_n-\gamma_n}}^{an^{1-\delta_n}} \frac{\theta_n}{n^{1-\delta_n}} \exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j > t) dt$$
$$\leq K(\varepsilon_2) \int_{an^{1-\delta_n-\gamma_n}}^{an^{1-\delta_n}} \frac{\theta_n}{n^{1-\delta_n}} \exp\left(g_n(t)\right) dt\ ,$$

where

$$g_n(t) = \frac{\theta_n t}{n^{1-\delta_n}} - c(1-\varepsilon_2)t^{\beta}.$$

Note that

$$\max_{an^{1-\delta_n-\gamma_n} \leq t \leq an^{1-\delta_n}} g_n(t) = g_n(an^{1-\delta_n}),$$

thus

$$\int_{an^{1-\delta_n-\gamma_n}}^{an^{1-\delta_n}} \frac{\theta_n}{n^{1-\delta_n}} \exp\left(g_n(t)\right) dt \leq an^{1-\delta_n} \exp\left(a\theta_n - c(1-\varepsilon_2)a^{\beta} n^{(1-\delta_n)\beta}\right)$$
$$= an^{1-\delta_n} \exp\left(a\Delta n^{\gamma_n} - c(1-\varepsilon_2)a^{\beta} n^{(1-\delta_n)\beta}\right).$$

Now, let us pick $D > 0$ and note that

$$\gamma_n = (1-\delta_n)\beta - D\delta_n < 1-\delta_n,$$

so we have that

$$an^{1-\delta_n} \exp\left(a\Delta n^{\gamma_n} - c(1-\varepsilon_2)a^{\beta} n^{(1-\delta_n)\beta}\right)$$
$$\leq an^{1-\delta_n} \exp\left(-n^{(1-\delta_n)\beta}[a^{\beta}c(1-\varepsilon_2) - a\Delta\exp(-D\varepsilon_0)]\right).$$

We can clearly select $D := D(\varepsilon_0, \Delta)$ such that

$$\exp(-D\varepsilon_0) = a^{\beta-1}c(1-2\varepsilon_2)/\Delta) \tag{22}$$

and obtain

$$\int_{an^{1-\delta_n-\gamma_n}}^{an^{1-\delta_n}} \frac{\theta_n}{n^{1-\delta_n}} \exp\left(\theta_n t/n^{1-\delta_n}\right) EZ_j^2 I(Z_j > t)dt$$
$$\leq K(\varepsilon_2)\, an^{1-\delta_n} \exp\left(-n^{(1-\delta_n)\beta}[\varepsilon_2 a^\beta c]\right).$$

Overall, then we conclude (redefining $K(\varepsilon_2)$)

$$E\exp\left(\theta_n Z_j/n^{1-\delta_n}\right) I(Z_j \leq an^{1-\delta_n})$$
$$= 1 + \frac{K(\varepsilon_2)}{n^{2(1-\delta_n)}} + O\left(\exp\left(-\widetilde{c}(\varepsilon_2)n^\beta\right)\right)$$

for some $\widetilde{c}(\varepsilon_2) > 0$ independent of the selection of the ε_j's, as long as $D(\Delta)$ is selected as indicated above. We then there exists $\widetilde{K}(\varepsilon_2)$ such that

$$(E\exp\left(\theta_n Z_0/n^{1-\delta_n}\right) I(Z_0 \leq an^{1-\delta_n}))$$
$$\times \left(E\exp\left(\theta_n Z_1/n^{1-\delta_n}\right) I(Z_1 \leq an^{1-\delta_n})\right)^n \exp\left(-\theta_n an^{\delta_n}\right).$$
$$\leq \exp(K(\varepsilon_2)\, n^{2\delta_n-1} + \widetilde{K}(\varepsilon_2)) \exp\left(-\Delta an^{\delta_n+\gamma_n}\right)$$
$$\leq \exp(K(\varepsilon_2)\, n^{2\delta_n-1} + \widetilde{K}(\varepsilon_2)) \exp\left(-\Delta an^{\beta+\delta_n(1-\beta-D)}\right)$$
$$= \exp(K(\varepsilon_2)\, n^{2\delta_n-1} + \widetilde{K}(\varepsilon_2)) \exp\left(-\Delta an^{\beta} \exp(\varepsilon_0(1-\beta)) \exp(-D\varepsilon_0)\right)$$
$$= \exp(K(\varepsilon_2)\, n^{2\delta_n-1} + \widetilde{K}(\varepsilon_2)) \exp(-a^\beta n^\beta c \exp(\varepsilon_0(1-\beta))(1-2\varepsilon_2)),$$

where in the last equality we used (22). We then obtain that

$$P(Z_0 + Z_1 + \cdots + Z_n > an, \cap_{j=0}^n \{Z_j \leq an^{1-\delta_n}\})$$
$$\leq \exp(K(\varepsilon_2)\, n^{2\delta_n-1} + \widetilde{K}(\varepsilon_2)) \exp(-a^\beta n^\beta c \exp(\varepsilon_0(1-\beta))(1-2\varepsilon_2)). \tag{23}$$

So, in summary, we have from (20)

$$\liminf_{n\to\infty} \frac{1}{n^\beta} \log P(Z_0 + Z_1 + \cdots + Z_n > an)$$
$$\geq \liminf_{n\to\infty} \frac{1}{n^\beta} \log P(Z_0 + Z_1 + \cdots + Z_n > an, \cup_{j=0}^n \{Z_j > an^{1-\delta_n}\})$$
$$\geq -c(a+\varepsilon_1)^\beta.$$

On the other hand, combining (19) and (23) we obtain

$$\liminf_{n\to\infty} \frac{1}{n^{\beta}} \log P\left(Z_0 + Z_1 + \cdots + Z_n > an\right)$$
$$\leq \max(-ca^{\beta}\exp\left(-\beta\varepsilon_0\right), -ca^{\beta}n^{\beta}\exp(\varepsilon_0(1-\beta))(1-2\varepsilon_2)).$$

Since $\varepsilon_0, \varepsilon_1$ and ε_3 are arbitrary we conclude the result. ■

Proof of Proposition 2. Note that for any $r \in (0,1)$

$$P_{\pi}\left(Y_0\left(p\right) > t\right) = P_{\pi}(\frac{1}{t}\int_0^{T_0} Q\left(\lfloor s \rfloor\right)^p ds > 1)$$
$$= P_{\pi}(\int_0^{T_0/t^r} (Q\left(\lfloor ut^{\xi}\rfloor\right)/t^{(1-\xi)/p})^p du > 1).$$

Pick $r = 1/(p+1)$ so that $r = (1-r)/p$. Let $x = t^r$ and recall that $\pi\left(k\right) = P\left(Q\left(\infty\right) = k\right) = \rho^k(1-\rho)$, so we can write

$$P_{\pi}\left(Y_0\left(p\right) > t\right) = P_{\pi}(\int_0^{T_0/x} (Q\left(\lfloor ux \rfloor\right)/x)^p du > 1)$$
$$= \sum_{k=0}^{\infty} \pi\left(k\right) P_k(\int_0^{T_0/x} (Q\left(\lfloor ux \rfloor\right)/x)^p du > 1)$$
$$= (1-\rho)\int_0^{\infty} \rho^{\lfloor s \rfloor} P_{\lfloor s \rfloor}(\int_0^{T_0/x} (Q\left(\lfloor ux \rfloor\right)/x)^p du > 1) ds$$
$$= (1-\rho)\int_0^{\infty} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(\int_0^{T_0/x} (Q\left(\lfloor ux \rfloor\right)/x)^p du > 1) dy.$$

Moreover, as mentioned earlier before we stated our current proposition,

$$\int_0^{\infty} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(\int_0^{T_0/x} (Q\left(\lfloor ux \rfloor\right)/x)^p du > 1) dy$$
$$= \int_0^{\infty} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du > 1) dy\ .$$

By the (functional) Law of Large Numbers, we have that conditional on $R\left(0\right) = \lfloor xy \rfloor$,

$$\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du$$
$$\longrightarrow \int_0^{y/\left|EW(1)\right|} (-u\left|EW\left(1\right)\right| + y)^p du = \int_0^y s^p ds = \frac{y^{p+1}}{p+1}.$$

So, if $y > (p+1)^{1/(p+1)} + \varepsilon$ for any $\varepsilon > 0$, we have that

$$\lim_{x\to\infty} P_{\lfloor xy \rfloor}(\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du > 1) = 1.$$

This implies that

$$\lim_{x\to\infty} \frac{1}{x} \log \int_{(p+1)^{1/(p+1)}}^{\infty} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du > 1) dy$$
$$= -(p+1)^{1/(p+1)} \log(1/\rho).$$

On the other hand, if $y \le (p+1)^{1/(p+1)} - \varepsilon$ we can apply large deviations theory. To do this let $\widehat{R}_x(\cdot)$ be the continuous piecewise linear approximation to $R(\lfloor \cdot x \rfloor)/x$ such that $x\widehat{R}_x(k/x) = R(k)$ for each integer $k \ge 0$. The process $\widehat{R}_x(\cdot)$ is an exponentially good approximation to $R(\lfloor \cdot x \rfloor)/x$, (see [4]). So the rate function (10) given by Mogulskii's theorem applies to $\widehat{R}_x(\cdot)$. Now, given a continuous function $\phi(\cdot)$ let $\mathcal{T}_-(\phi) = \inf\{s > 0 : \phi(s) < 0\}$ and note that the mapping $F(\cdot)$ defined as

$$F(\phi) = \int_0^{\mathcal{T}_-(\phi)} \phi(s)^p \, ds$$

is a continuous mapping under the uniform topology on compact sets. So, by the contraction principle it follows easily that

$$\frac{1}{x} \log \int_0^{(p+1)^{1/(p+1)}} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(F(\widehat{R}_x) > 1) dy \to -C(p), \tag{24}$$

where $C(p)$ is the solution of the variational problem stated in (11). Clearly, $\mathcal{T}_-(\widehat{R}_x) \le \mathcal{T}/x$ so we have that

$$P_{\lfloor xy \rfloor}(\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du > 1) \le P_{\lfloor xy \rfloor}(F(\widehat{R}_x) > 1).$$

This implies that

$$\overline{\lim}_{x\to\infty} \frac{1}{x} \log \int_0^{(p+1)^{1/(p+1)}} x\rho^{\lfloor xy \rfloor} P_{\lfloor xy \rfloor}(\int_0^{\mathcal{T}/x} (R(\lfloor ux \rfloor)/x)^p du > 1) dy \le -C(p).$$

The lower bound is obtained simply by tracking a continuous path that is arbitrarily close the solution of the calculus of variations problem.

■

References

[1] S. Asmussen. *Applied Probability and Queues*. Springer-Verlag, New York, 2003.

[2] J. Blanchet. Optimal sampling of overflow paths in jackson networks. *Preprint*, 2010.

[3] A. A. Borovkov. Estimates for the distribution of sums and maxima of sums of random variables without the cramer condition. *Siberian Math. J.*, 41:811–848, 2000.

[4] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*. Springer, New York, second edition, 1998.

[5] D. Denisov, A. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.*, 36:1946–1991, 2008.

[6] M.D. Donsker and S.R.S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28:1–47; ibid. **28** (1975), 279–301, 1975.

[7] K. R. Duffy and S. P. Meyn. Most likely paths to error when estimating the mean of a reflected random walk. *Performance Evaluation*, 67(12):1290–1303, 2010.

[8] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, New York, 1997.

[9] M.I. Freidlin and A .D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 1984. Series in Comprehensive Studies in Mathematics.

[10] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13:304–362, 2003.

[11] S. P. Meyn. Large deviation asymptotics and control variates for simulating large functions. *Ann. Appl. Probab.*, 16(1):310–339, 2006.

[12] S. P. Meyn. *Control Techniques for Complex Networks.* Cambridge University Press, Cambridge, 2007.

[13] H. Miller. A convexity property in the theory of random variables defined on a finite markov chain. *Ann. of Math. Stat.*, 32:1260–1270, 1961.

[14] P. Ney and E. Nummelin. Markov additive processes II. Large deviations. *Ann. Probab.*, 15(2):593–609, 1987.

[15] L. V. Rozovskii. Probabilities of large deviations of sums of independent random variables. *Theory Probab. Appl.*, 34:625–644, 1990.