ASYMPTOTICS OF THE AREA UNDER THE GRAPH OF A LÉVY-DRIVEN WORKLOAD PROCESS

J. Blanchet and M. Mandjes

Abstract. Let \((Q_t)_{t \in \mathbb{R}}\) be the stationary workload process of a Lévy-driven queue, where the driving Lévy process is light-tailed. For various functions \(T(u)\), we analyze
\[
P\left( \int_0^T Q_s ds > u \right)
\]
for \(u\) large. For \(T(u) = o(\sqrt{u})\) the asymptotics resemble those of the steady-state workload being larger than \(u/T(u)\). If \(T(u)\) is proportional to \(\sqrt{u}\) they look like \(e^{-\alpha \sqrt{u}}\) for some \(\alpha > 0\). Interestingly, the asymptotics are still valid when \(\sqrt{u} = o(T(u))\), \(T(u) = o(u)\), and \(T(u) = \beta u\) for \(\beta\) suitably small.

1. Introduction

Let \((X_t)_{t \in \mathbb{R}}\) be a two sided (one-dimensional) Lévy process with \(X_0 = 0\). It is commonly known that under the stability condition \(\mathbb{E}X_1 < 0\) the stationary workload process, given by \(Q_t := \sup_{s \in (-\infty,t]} (X_t - X_s)\), is well-defined. It is clear that
\[
\frac{1}{T} \int_0^T Q_s ds
\]
converges to the steady-state mean workload \(\mathbb{E}Q_0\) as \(T \to \infty\), as a direct consequence of the ergodic theorem.

However, so far hardly any explicit results are available on the random variable \((\mathbb{I})\). The objective of the present paper is to shed light on this, by studying the large-deviations probabilities
\[
\pi_{T(u)}(u) := P\left( \int_0^T Q_s ds > u \right)
\]
for \(u\) large, and various types of functions \(T(u)\); the workload is assumed to be in stationarity at time zero.

The relevance of this study is that it exhibits an important example—one could even argue that this is actually the most fundamental example that would come to the mind to a queueing theorist—for which the standard Donsker-Varadhan large-deviations theory is not directly applicable, even under natural assumptions. For instance, as we shall see, if \(T(u) = u\beta\), with \(\beta < 1/\mathbb{E}Q_0\), then the asymptotics for \(\pi_{T(u)}(u)\) are subexponential in \(u\).

Date: February 9, 2013.

Key words and phrases. Queues, workload process, area, large deviations, Lévy processes.
despite assuming that $(X_t)_{t \in \mathbb{R}}$ is a well-behaved, light-tailed Lévy process. In this paper we study the impact of the function $T(\cdot)$ on the tail asymptotics of $(1)$. More specifically, under natural large deviations conditions, we obtain the following results:

- For $T(u) = o(\sqrt{u})$ the asymptotics resemble those of the steady-state workload being larger than $u/T(u)$:

$$\log \pi_{T(u)}(u) \sim \log \mathbb{P}(Q_0 > \frac{u}{T(u)}).$$

Intuitively this result means that, in order to ensure that the area is larger than $u$, essentially it is just required that $Q_0$ is larger than $u/T(u)$.

- For $T(u)$ of the form $T(\sqrt{u})$, it is shown that

$$\lim_{u \to \infty} \frac{1}{\sqrt{u}} \log \pi_{T(u)}(u) = -\alpha,$$

for a constant $\alpha > 0$ that is identified explicitly. For $T$ small (that is, $T$ below an explicitly given threshold), the most likely way in which the rare event happens is such that, with overwhelming probability, the buffer never idles in $[0, T\sqrt{u}]$; for $T$ above the threshold the most likely path consists essentially of a single ‘big’ busy period.

- Finally it is shown that $(2)$ remains valid in case $\sqrt{u} = o(T(u))$ and $T(u) = o(u)$, and also if $T(u) = \beta u$ for $\beta < 1/\mathbb{E}Q_0$.

The special case of $(X_t)_{t \in \mathbb{R}}$ corresponding to Brownian motion was already covered in $[1]$. Many steps in the analysis presented in $[1]$ used explicit properties of Brownian motion that are not available for light-tailed Lévy processes. Indeed, one of the challenges of the present paper was to find their counterparts for our more general light-tailed Lévy setting. In addition, the case of $T(u) = \beta u$ with $\beta < 1/\mathbb{E}Q_0$ was not covered in $[1]$. Also the paper $[5]$ is strongly related to ours. There $\pi_{T(u)}(u)$ is analyzed for the number of customers in the M/M/1 queue, with $T(u) = \beta u$ with $\beta < 1/\mathbb{E}Q_0$. At the methodological level, the analysis presented in this paper borrows elements from both $[1]$ and $[5]$.

Our paper is organized as follows. In Section 2 the main objective is to find the asymptotics for the auxiliary object

$$\varrho_t := \mathbb{P}\left(\frac{1}{\sqrt{t}} \int_0^t X_s ds > a\right);$$

this result, which is is extensively used later in the paper, relies precisely on Assumption1.

Then Section 3 covers the case that $T(u)$ is small relative to $\sqrt{u}$; the analysis relies on straightforward bounds in combination with the classical bound $\mathbb{P}(Q_0 > u) \leq e^{-\kappa u}$ $[3]$. In Section 4 $T(u)$ is taken proportional to $\sqrt{u}$. Under the assumption that a sample-path large deviations principle is valid, $[2]$ is established. Section 5 covers the cases (i) $\sqrt{u} = o(T(u))$ and $T(u) = o(u)$, and (ii) $T(u) = \beta u$ with $\beta < 1/\mathbb{E}Q_0$. In $[5]$ it was shown that the probability that the sum of $\gamma u$ Weibullian random variables, each of them behaving
as $e^{-\alpha\sqrt{a}}$, exceeding $u$, essentially behaves as a single of those Weibullians exceeding $u$ (in terms of logarithmic asymptotics). Relying on this property it is shown that (2) is valid in this case as well. Finally, Section 6 we discuss how to relax some of the simplifying assumptions that we imposed, and how to obtain results assuming non-stationary initial conditions. Section 7 provides some technical large deviations results.

2. ASYMMETRY OF THE INTEGRAL OF A LÉVY PROCESS

The main objective of this section is to study large deviations asymptotics for the auxiliary quantity

$$\varrho_t := \mathbb{P}\left( \frac{1}{t^2} \int_0^t X_s ds > a \right).$$

In our future analysis of $\pi_T(u)$, which is the main goal of the paper, only logarithmic results for $\varrho_t$ might suffice; however, because we believe that this quantity is of independent interest and because a sharp analysis is not difficult to perform, we provide exact asymptotics.

First, define $\phi(\vartheta) := \log \mathbb{E}e^{\vartheta X_1}$ and let $Z(t) := \int_0^t X_s ds$. Introducing the notation $\eta_t(\vartheta) := \mathbb{E}e^{\vartheta Z(t)}$ and applying integration by parts, we obtain

$$\int_0^t X_s ds = tX_t - \int_0^t s dX_s = \int_0^t (t - s) dX_s,$$

to conclude that

$$\eta_t(\vartheta) = \exp \left( \int_0^t \phi \left( \vartheta \left( \frac{t - s}{t} \right) \right) ds \right) = \exp \left( t \int_0^1 \phi(\vartheta u) du \right).$$

Consequently, we have that

$$\chi_t(\vartheta) := \frac{1}{t} \log \eta_t(\vartheta) = \int_0^1 \phi(\vartheta u) du.$$

The quantities $\chi_t(\vartheta)$ and $\eta_t(\vartheta)$ play an important role in the characterization of the exact asymptotics of $\varrho_t$. In order to develop such asymptotics, we shall impose throughout the rest of the paper the following assumption, which is fairly standard when developing large deviation estimates.

**Assumption 1.** (Steepness to the right): If $\phi(\vartheta) = \log(\mathbb{E}e^{\vartheta X_1})$, then for every $a > \mathbb{E}X_1$ there exists $\vartheta^* > 0$ such that $\phi'(\vartheta^*) = a$.

Using expression (3) we can obtain logarithmic asymptotics (via an application, for instance, of the Gärtner-Ellis theorem, see [8]) leading to

$$\lim_{t \to \infty} \frac{1}{t} \log \varrho_t = -\sup_{\vartheta \geq 0} \left( a\vartheta - \int_0^1 \phi(\vartheta x) dx \right) =: -J(a).$$

The convexity of $\phi(\cdot)$ implies the convexity of $\vartheta \mapsto \int_0^1 \phi(\vartheta x) dx$, which in turn implies (together with the fact that $\phi(0) = 0$, and that $a > \phi'(0)/2$) that the supremum in the previous display is the same if one optimizes over $\vartheta \in \mathbb{R}$. Moreover, it also follows that conditions for local optimality imply global optimality in the optimization problem.
underlying the definition of $J(a)$. The next lemma, therefore, shows that there is a unique optimizer $\vartheta^*$ to the previous optimization problem.

**Lemma 1.** For every $a > \mathbb{E}Y/2$, there exists $\vartheta^* > 0$ such that

$$a = \int_0^1 \phi'(\vartheta^* x) x \, dx.$$

**Proof:** Because of the monotone convergence theorem, it follows that for $\vartheta \in (0, \theta^*)$,

$$\frac{d}{d\vartheta} \int_0^1 \phi(\vartheta x) \, dx = \int_0^1 \phi'(\vartheta x) x \, dx,$$

where $\phi'(\cdot)$ is the derivative of $\phi(\cdot)$. In turn, we have that

$$\int_0^1 \phi'(\vartheta x) \, dx = \frac{1}{\vartheta^2} \int_0^\vartheta \phi'(y) y \, dy.$$

We must show that there exists a unique solution to the equation

$$(5) \quad \vartheta^2 \left( a - \phi'(0)/2 \right) = \int_0^\vartheta (\phi'(y) - \phi'(0)) y \, dy.$$

By the strict convexity of $\phi$, which follows because $Y$ is a non-degenerate random variable in view of Assumption 1, we have that $\phi'(y) - \phi'(0) > 0$ if $y > 0$. Moreover, both the right-hand side and the left-hand side of Eqn. (5) are convex functions of $\vartheta$. The derivative of the left-hand side is larger than the derivative of the right-hand side for values of $\vartheta$ that are sufficiently close to zero, but because of Assumption 1 eventually the derivative of the right-hand side increases superlinearly in $\vartheta$. So, eventually the right hand side overtakes the left-hand side. By continuity, thus, there exists a solution to Eqn. (5). The solution must be unique because the optimization problem underlying the definition of $J(a)$ is a strictly concave program.

Now we are ready to sharpen the large deviations result (4).

**Theorem 1.** Suppose, in addition to Assumption 1 that $\phi(\theta) < \infty$ for $\theta$ in a neighborhood of the origin. Define

$$\sigma^2 := \int_0^1 \phi''(\vartheta^* x) x^2 \, dx.$$

Then, as $t \to \infty$,

$$\varrho_t := \mathbb{P}(Z(t) > at) \sim \frac{1}{\sqrt{t}} \frac{1}{\sqrt{2\pi \vartheta^* \sigma}} e^{-tJ(a)}.$$

**Proof of Thm. 1.** According to [6, Thm. 3.3] we need to verify the following properties. First, it should be checked that $\chi_t(\cdot)$ is analytic, for every $t > 0$, in a neighborhood of the origin. This property is immediate from the fact that $\chi_t(\cdot) = \int_0^1 \phi(\cdot u) \, du$ is independent of $t$, and $\phi(\cdot)$ is clearly analytic around the origin since we are assuming the existence of exponential moments of $X_1$ in a neighborhood of the origin. Second, we need to show
that $\chi''(\vartheta) > 0$. Applying ‘dominated convergence’ (which is justified due to Assumption 1) we obtain that
\[
\sigma^2 = \chi''(\vartheta^*) = \int_0^1 \phi''(\vartheta^* x) x^2 \, dx > 0;
\]
the inequality follows because $X_1$ cannot be a degenerate random variable (i.e., a deterministic quantity), once again, due to Assumption 1. Finally, the third and last property that we need to verify is that that there exists $\delta_1$ such that for any $\delta_0, \lambda > 0$, with the property that $0 < \delta_0 < \delta_1 < \lambda$,
\[
\lim_{t \to \infty} \sup_{\delta_0 \leq |\theta| \leq \lambda} \frac{|\eta_t(\vartheta^* + i \theta)|}{|\eta_t(\vartheta^*)|} = o \left( \frac{1}{\sqrt{t}} \right)
\]
as $t \to \infty$, where $|z|$ denotes the modulus of a complex number $z$ (so if $z$ is a real number the definition corresponds to the absolute value). To this end, first note that $\eta_t(\vartheta^* + i \cdot) / \eta_t(\vartheta^*)$ is simply the characteristic function of a member of the natural exponential family of $Z(t)$, corresponding to the natural parameter $\vartheta^*$. So, it suffices to show that for any random variable such as $Z(t)$ one can find $\delta_1 > 0$ such that for any $\delta_0, \lambda > 0$ satisfying $0 < \delta_0 < \delta_1 < \lambda$,
\[
(6) \quad \sup_{\delta_0 \leq |\theta| \leq \lambda} \left| \exp \left( \int_0^1 \phi(i \theta x) \, dx \right) \right| < 1.
\]
It is now noted that this property holds, as can be seen as follows. Suppose first that $X_1$ has a lattice distribution with span $h$. If $0 < \delta_0 < \delta_1 < \min (\pi / h, 1)$ then we have that
\[
\xi(\delta_0, \delta_1) := \sup_{\delta_0 \leq |\theta| \leq \delta_1} \left| \exp \left( \phi(i \theta) \right) \right| < 1.
\]
Now, observe that for any $v < w$, the quantity $\exp(\int_v^w \phi(i \theta x) \, dx)$ can be interpreted as the characteristic function of the random variable $\int_v^w s \, dX_s$, and hence such quantity has modulus less than or equal to unity. Consequently, if $0 < v < w < 1$, then
\[
(7) \quad \left| \exp \left( \int_0^1 \phi(i \theta x) \, dx \right) \right| \leq \exp \left( \int_v^w \phi(i \theta x) \, dx \right).
\]
Now, given $\delta_0, \lambda > 0$, select $w < \delta_1 / \lambda$ and $v = \delta_0$, to conclude that if $|\theta| \in [\delta_0, \lambda]$, then
\[
(8) \quad \left| \exp \left( \int_v^w \phi(i \theta x) \, dx \right) \right| = \lim_{n \to \infty} \prod_{k=0}^n \left| \exp \left( \phi \left( i \theta \left( v + \left( w - v \right) \frac{k}{n} \right) \right) \frac{w - v}{n} \right) \right| \leq \xi^{w-v} < 1.
\]
Inequalities (7) and (8) together imply (6) and thus the result. The case in which $X_1$ is even easier as in this case one can select $\delta_1 = 1$. \qed
3. SHORT TIMESCALE

From now on we consider the queueing model described in the introduction; in this section we concentrate on the short timescale regime in which $T(u) = o(\sqrt{u})$. Throughout the rest of the paper, in addition to Assumption 1 we shall impose the following standard large deviations assumptions.

**Assumption 2.** *(Cramér condition)* There is $\kappa > 0$ such that $\mathbb{E} \exp (\kappa X_1) = 1$.

Under Assumption 2 the celebrated Cramér-Lundberg asymptotics hold, meaning that the tail probability $\mathbb{P}(Q_0 > u)$ decays exponentially in $u$. More precisely, as follows from [3], with the $\kappa$ introduced in Assumption 2 the asymptotics take the form, as $u \to \infty$,

$$\log \mathbb{P}(Q_0 > u) = -\kappa u (1 + o(1)).$$

In addition, we impose the following simplifying assumption.

**Assumption 3.** There exists $\delta < 0$ such that $\mathbb{E} \exp (\delta |X_1|) < \infty$.

As we shall discuss in Section 6, the results in this paper can be derived only under Assumptions 1 and 2. Assumption 3 allows us to take advantage of sample-path large deviations, a concept that we shall review now.

With the local rate function given by $I(x) := \sup_{\vartheta} (\vartheta x - \phi(\vartheta))$, we define the sample-path rate function $\mathbb{I}(\cdot)$, for $T > 0$, by

$$\mathbb{I}(f) = \int_0^T I(f'(t)) dt$$

for any absolutely continuous $f$, and $\infty$ else. A sample-path large deviations principle then says that for any closed $A$

$$\limsup_{u \to \infty} \frac{1}{u} \log \mathbb{P} \left( \frac{X_u}{u} \in A \right) \leq -\inf_{f \in A} \mathbb{I}(f),$$

and for any open $B$

$$\liminf_{u \to \infty} \frac{1}{u} \log \mathbb{P} \left( \frac{X_u}{u} \in B \right) \geq -\inf_{f \in B} \mathbb{I}(f).$$

To determine whether a given set is open or closed we use the norm

$$\|f\| := \sup_{s \in (-\infty, \infty)} \frac{|f(s)|}{1 + |s|}.$$

We state the following technical result whose proof is given in the appendix at the end of this paper.

**Lemma 2.** The sample-path large deviations principle is valid under the norm $\|f\|$.

It will also be useful to keep in mind the representation

$$Q_t = \max \left( Q_0 + X_t, \max_{0 \leq s \leq t} (X_t - X_s) \right).$$
Recall that $\kappa > 0$ solves $E e^{\kappa X_1} = 1$, and, as earlier,
\[ J(a) := \sup_{\vartheta > 0} \left( a \vartheta - \int_0^1 \phi(\vartheta x)dx \right). \]
The main result of this section states that if $T(u) = o(\sqrt{u})$, the probabilities $\pi_T(u)$ and $P(Q_0 > u/T(u))$ are equivalent on a logarithmic scale.

**Theorem 2.** If $T(u) = o(\sqrt{u})$, then
\[ \lim_{u \to \infty} \frac{T(u)}{u} \log \pi_T(u) = -\kappa. \]

**Lower bound:** Using the first term in the right-hand side of (9),
\[ \pi_T(u) \geq P \left( Q_0 T(u) + \int_0^{T(u)} X_s ds > u \right). \]
Under the light-tailed conditions, we have
\[ \lim_{u \to \infty} \frac{T(u)}{u} \log P \left( \exists t \in [0,1] : Q_t \geq u \right) = -\kappa, \]
where $\kappa > 0$ solves $\phi(\kappa) = 0$. Also, using Chernoff’s bound and our expression for $J(\cdot)$ derived in the previous section we obtain
\[ P \left( \int_0^{T(u)} X_s ds > u \right) \leq \exp \left( -T(u) J \left( \frac{u}{(T(u))^2} \right) \right). \]
Hence, to prove the lower bound, it suffices to prove that for $u$ sufficiently large
\[ \kappa \frac{u}{(T(u))^2} < J \left( \frac{u}{(T(u))^2} \right). \]
This follows from the strict convexity of $J(\cdot)$ in conjunction with $u/(T(u))^2 \to \infty$ as $u \to \infty$. This completes the proof of the lower bound.

**Upper bound:** Note that, due to the union bound and the stationarity of the workload process,
\[ \pi_T(u) \leq P \left( \int_0^1 Q_s ds + \int_1^2 Q_s ds + \ldots + \int_{T(u)}^{T(u)+1} Q_s ds > u \right) \]
\[ \leq (T(u) + 1) P \left( \exists t \in [0,1] : Q_t \geq \frac{u}{T(u)} \right), \]
so that it suffices to prove that
\[ \limsup_{u \to \infty} \frac{T(u)}{u} \log P \left( \exists t \in [0,1] : Q_t \geq \frac{u}{T(u)} \right) \leq -\kappa. \]
Now, from (9) we obtain that
\[ P \left( \exists t \in [0,1] : Q_t \geq \frac{u}{T(u)} \right) \leq P \left( \exists t \in [0,1] : Q_0 + X_t \geq \frac{u}{T(u)} \right) \]
\[ + \sup_{0 \leq s \leq t} (X_t - X_s) \geq \frac{u}{T(u)}. \]
We now analyze both probabilities in the right-hand side of the previous display in detail.

- Note that for each \( \varepsilon > 0 \) we have that
  \[
  P \left( \exists t \in [0, 1] : Q_0 + X_t \geq \frac{u}{T(u)} \right) 
  \leq P \left( Q_0 \geq (1 - \varepsilon) \frac{u}{T(u)} \right) + P \left( \max_{0 \leq t \leq 1} X_t \geq \varepsilon \frac{u}{T(u)} \right).
  \]

We know that, by Assumption 2,
\[
\lim_{u \to \infty} \frac{T(u)}{u} \log P \left( Q_0 \geq (1 - \varepsilon) \frac{u}{T(u)} \right) = -\kappa (1 - \varepsilon). \tag{10}
\]

In addition, with \( n(u) := u/T(u) \) and \( \delta > 1/n(u) \),
\[
P \left( \max_{0 \leq t \leq 1} X_t \geq \varepsilon \frac{u}{T(u)} \right) \leq P \left( \max_{0 \leq t \leq \delta} \frac{X_{n(u)t}}{n(u)} \geq \varepsilon \right). \]

Because of Lemma 2,
\[
\limsup_{u \to \infty} \frac{1}{n(u)} \log P \left( \max_{0 \leq t \leq \delta} \frac{X_{n(u)t}}{n(u)} \geq \varepsilon \right) = -\delta I \left( \frac{\varepsilon}{\delta} \right).
\]

By the strict convexity of \( I(\cdot) \), we have that \( I(x/\delta) \geq \kappa x/\delta \) for \( \delta \) sufficiently small. Combining this with (10), we conclude that
\[
\limsup_{u \to \infty} \frac{T(u)}{u} \log P \left( \exists t \in [0, 1] : Q_0 + X_t \geq \frac{u}{T(u)} \right) \leq -\kappa.
\]

- On the other hand, with \( n(u) = u/T(u) \) for any \( \delta > 1/n(u) \), we note that
  \[
P \left( \exists t \in [0, 1] : \max_{0 \leq s \leq t} (X_t - X_s) \geq \frac{u}{T(u)} \right)
  = P \left( \max_{t \in [0,1]} \max_{0 \leq s \leq t} (X_t - X_s) \geq \frac{u}{T(u)} \right)
  \leq P \left( \max_{0 \leq t \leq \delta} \max_{0 \leq s \leq t} \frac{X_{n(u)t} - X_{n(u)s}}{n(u)} \geq 1 \right).
  \]

By Lemma 2,
\[
\limsup_{u \to \infty} \frac{1}{n(u)} \log P \left( \max_{0 \leq t \leq \delta} \max_{0 \leq s \leq t} \frac{X_{n(u)t} - X_{n(u)s}}{n(u)} \geq 1 \right) = -\delta I \left( \frac{1}{\delta} \right).
\]

As \( \delta > 0 \) can be chosen arbitrarily small, the convexity of \( I(\cdot) \) entails that \( \delta I(1/\delta) \) can be made arbitrarily large and we conclude the proof.

4. INTERMEDIATE TIMESCALE

In this section we analyze
\[
\lim_{u \to \infty} \frac{1}{\sqrt{u}} \log P \left( \int_0^{Tu} Q_s ds \geq Mu \right) = \lim_{u \to \infty} \frac{1}{u} \log P \left( \int_0^{Tu} Q_s ds \geq Mu^2 \right).
\]
The probability in the right-hand side of the previous display can be rewritten immediately in terms of an expressions that can be analyzed using a sample-path large deviations principle for Lévy processes. Indeed, observe that

\[
\mathbb{P}\left( \int_0^{Tu} Q_s ds \geq Mu^2 \right) = \mathbb{P}\left( \int_0^{Tu} \sup_{s \leq r \leq T} (X_r - X_s) ds \geq Mu^2 \right) = \mathbb{P}\left( \int_0^{Tu} \sup_{r \leq Tu} (X_r - X_s) dt \geq Mu \right) = \mathbb{P}\left( \frac{1}{u} \int_0^{T} \sup_{0 \leq s \leq t} (X_t - X_s) dt \geq M \right) = \mathbb{P}\left( \frac{1}{u} \int_0^{T} Q_{tu} dt \geq M \right) = \mathbb{P}\left( X_{\cdot} u \in A \right),
\]

with

\[
A := \left\{ f : \int_0^{T} \sup_{0 \leq s \leq t} (f(t) - f(s)) dt \geq M \right\}.
\]

We can now state the main result of this section.

**Theorem 3.**

\[
\lim_{u \to \infty} \frac{1}{u} \log \mathbb{P}\left( \int_0^{Tu} Q_s ds \geq Mu^2 \right) = -\inf_{a \geq 0} \inf_{s \in [0,T]} \left( \kappa a + sJ\left( \frac{M - as}{s^2} \right) \right).
\]

To prove this theorem, we first need to show an auxiliary result. To this end, define

\[ p_u(T, M, a) := \mathbb{P}\left( \int_0^{Tu} Q_s ds \geq Mu \mid Q_0 = au \right). \]

**Lemma 3.**

\[
\lim_{u \to \infty} \frac{1}{u} \log p_u(T, M, a) \leq -\inf_{s \in [0,T]} sJ\left( \frac{M - as}{s^2} \right).
\]

**Proof:** The proof of this lemma is very similar to the proof of the Brownian case in [1], and we therefore just sketch the most important steps. First it is observed that \( A \) (as defined in (11)) is closed; see the proof of the second part in the appendix of [1]. As a result,

\[
\lim_{u \to \infty} \frac{1}{u} \log p_u(T, M, a) \leq -\inf_{f \in A} I(f).
\]

Now take an arbitrary path \( f \) in \( A \) (where \( f \) is absolutely continuous), we construct \( \bar{f} \) as follows. Define \( q(f)(t) := \sup_{0 \leq s \leq f(t)} (f(t) - f(s)) \); \( f \) represents a trajectory of the input to the queue in fluid scale. We have subintervals of \([0, T]\) in which the queue is empty (idle periods), and subintervals in which it is non-empty. Clearly there are at most countably many idle periods because the total length of time the queue is empty is less than \( T \). Now, permute these such that all busy periods are moved to the front, and the idle periods to the back. For example, if \( f(s) = \min(s, 2-s)^+ + \min(s-4, 6-s)^+ \), then \( \bar{f}(s) = \min(s, 2-s)^+ + \min(s-2, 4-s)^+ \). The corresponding transformed input path, which we call \( \bar{f} \), has the property that \( I(f) = I(\bar{f}) \); to see this, the essential property here is that \( (X_t)_{t \in \mathbb{R}} \)
has stationary and independent increments, which yields that the large deviations rate function of a trajectory \( f \) just depends on the derivative of the path.

As a result,

\[
\inf_{f \in A} I(f) = \inf_{f \in \bar{M}} I(f), \quad \text{with} \quad \bar{M} := \left\{ f : \exists s \in [0, T] : \int_0^s f(r)dr \geq M - as \right\}.
\]

Evidently,

\[
\inf_{f \in \bar{M}} I(f) = \inf_{s \in [0, T]} \inf_{f \in \bar{M}_s} I(f), \quad \text{with} \quad \bar{M}_s := \left\{ f : \int_0^s f(r)dr \geq M - as \right\}.
\]

But, due to the sample-path large deviations principle,

\[
\inf_{f \in \bar{M}_s} I(f) = \lim_{u \to \infty} \frac{1}{u} \log \mathbb{P} \left( \int_0^s X_{tu}dt \geq (M - as)u \right) = -sJ \left( \frac{M - as}{s^2} \right),
\]

where the last equality is due to Thm. 1.

**Proof of Thm. 3** The proof consists of two bounds. The lower bound is straightforward, whereas the upper bound relies on Lemma 3.

**Lower bound:** The starting point is, for any \( s \in [0, T] \),

\[
\mathbb{P} \left( \frac{1}{u} \int_0^T Q_{tu}dt \geq M \right) \geq \mathbb{P} \left( \frac{1}{u} \int_0^s Q_{tu}dt \geq M \right) \geq \mathbb{P} \left( Q_0s + \int_0^s X_{tu}dt \geq Mu \right).
\]

Evidently, this majorizes, for any \( a \geq 0 \),

\[
\mathbb{P} (Q_0 \in [a, a + \varepsilon)u) \mathbb{P} \left( \int_0^s X_{tu}dt \geq (M - as)u \right).
\]

The exact asymptotics [3] of the first factor look like \( Ce^{-\kappa au}(1 - e^{-\kappa \varepsilon u}) \), for some \( C > 0 \), which has logarithmic asymptotics \( -\kappa a \). The second factor obeys, due to Thm. 1,

\[
\lim_{u \to \infty} \frac{1}{u} \log \mathbb{P} \left( \int_0^s X_{tu}dt \geq (M - as)u \right) = -sJ \left( \frac{M - as}{s^2} \right).
\]

As the lower bounds hold for any \( s \in [0, T] \) and \( a \geq 0 \), we obtain

\[
\liminf_{u \to \infty} \frac{1}{u} \log \mathbb{P} \left( \int_0^{Tu} Q_sds \geq Mu^2 \right) \geq -\inf_{a \geq 0} \inf_{s \in [0, T]} \left( \kappa a + sJ \left( \frac{M - as}{s^2} \right) \right).
\]
Upper bound: Clearly,
\[ P \left( \int_0^T u Q_s ds \geq Mu^2 \right) \]
\[ \leq \sum_{k=0}^{\infty} P(Q_0 \in [k\varepsilon, (k+1)\varepsilon) u)p_u(T, M, (k+1)\varepsilon) \]
\[ \leq \sum_{k=0}^{N-1} P(Q_0 \geq k\varepsilon u)p_u(T, M, (k+1)\varepsilon) + \sum_{k=N}^{\infty} P(Q_0 \geq k\varepsilon u) \]
\[ \leq \sum_{k=0}^{N-1} e^{-\kappa k\varepsilon u} p_u(T, M, (k+1)\varepsilon) + \sum_{k=N}^{\infty} e^{-\kappa \varepsilon u} \frac{1}{1 - e^{-\kappa \varepsilon u}}. \]

Applying [8, Lemma 1.2.15], it follows that the decay rate of interest is bounded from above by
\[ \max \left\{ \max_{k=0, \ldots, N-1} \left( \lim_{u \to \infty} \frac{1}{u} \log p_u(T, M, (k+1)\varepsilon) \right) - \kappa \cdot k\varepsilon, -\kappa \cdot N\varepsilon \right\}. \]

As a result of Lemma 3, we can bound the above decay further by
\[ \max \left\{ \max_{k=0, \ldots, N-1} \left( -\inf_{s \in [0, T]} sJ(M - (k+1)\varepsilon s) \frac{s}{s^2} \right) - \kappa \cdot k\varepsilon, -\kappa \cdot N\varepsilon \right\} \]
\[ \leq -\min \left\{ \inf_{a \geq 0} \inf_{s \in [0, T]} sJ(M - as) \frac{s}{s^2} + \kappa a - \kappa \varepsilon, \kappa \cdot N\varepsilon \right\}. \]

The upper bound then follows after letting \( N \uparrow \infty \), and then \( \varepsilon \downarrow 0 \). \qed

Remark 1. The decay rate
\[ \inf_{a \geq 0} \inf_{s \in [0, T]} \left( \kappa a + sJ(M - as) \frac{s}{s^2} \right) \]
can be analyzed in more detail. Define \( T^* \) as the solution to
\[ \kappa = J' \left( \frac{M}{T^2} \right), \]
which exists because \( J'(\cdot) \) is increasing (following from the fact that \( J(\cdot) \) is strictly convex by Assumption 2), and continuous
\[ J'(0) := \arg \inf_{\theta \geq 0} \int_0^1 \phi(\theta x) dx < \kappa. \]

Elementary analysis yields that for \( T \in [0, T^*] \), the infimum over \( s \) is attained at \( s^* = T \). Then the optimal \( a \) follows from optimizing
\[ \inf_{a \geq 0} \left( \kappa a + T J \left( \frac{M - aT}{T^2} \right) \right). \]
which is a convex function in $a$. It is seen that, due to the inequality $T < T^\star$, the maximum is attained for a positive $a^\star$; this follows from the definition of $T^\star$. The intuition behind this solution is that the most likely path is such that there is, most likely, already a positive buffer content at time 0, and the buffer remains non-idle during the entire interval $[0, T]$.

For $T \geq T^\star$, the optimum is achieved at $a^\star = 0$ and $T = T^\star$. In this case, with overwhelming probability, the buffer content starts off at a value near to 0 at time 0, to become idle again at roughly $T^\star$.

### 5. LONG TIME-SCALE

In this section we prove the following result.

**Theorem 4.** If either (i) $\sqrt{u} = o(T(u))$ and $T(u) = o(u)$, or (ii) $T(u) = \beta u$ for $\beta < 1/EQ$, then

$$
\lim_{u \to \infty} \frac{1}{\sqrt{u}} \log P \left( \int_0^{T(u)} Q_r \, dr > u \right) = -T^\star J \left( \frac{1}{(T^\star)^2} \right).
$$

**Proof of Thm. 4.** We argue the lower and upper bound separately.

**Lower bound:** The lower bound is trivial, as we have that

$$
P \left( \int_0^{T(u)} Q_r \, dr > u \right) \geq P \left( \int_0^{T\sqrt{u}} Q_r \, dr > u \right);
$$

applying the results of Section 4 yields the claim.

**Upper bound:** The upper bound is more involved, but can be established by applying techniques developed in [1] and [5]. Let us first consider case (i), that is, $\sqrt{u} = o(T(u))$ and $T(u) = o(u)$. Replace the process $Q_r$ by a process $\bar{Q}_r$, as follows. Start in stationarity, and let $\bar{Q}_r$ equal $Q_r$ until the buffer idles, say at time $\tau$. Then let $\bar{Q}_\tau$ be sampled from the stationary distribution, and let it follow the workload process from then on; again, as soon as it hits zero, it is replaced by a sample from the stationary distribution, etc. Clearly, $\bar{Q}_r \geq Q_r$, and with $N(u)$ the number of these i.i.d. cycles in $T(u)$, we have that

$$
\int_0^{T(u)} Q_r \, dr \leq \sum_{i=1}^{N(u)} H_i,
$$

with the $H_i$ distributed as a random variable $H$ such that

$$
H \overset{d}{=} \int_0^{T} Q_r \, dr.
$$

With the techniques used in the proof of Lemma 3, we can prove that

$$
\lim_{u \to \infty} \frac{1}{\sqrt{u}} \log P \left( H > Mu \right) \leq -T^\star J \left( \frac{M}{(T^\star)^2} \right).
$$

Similar to the upper bound in the proof of [1] Thm. 3, we have, with $K := 2/E\tau$,

$$
P \left( \sum_{i=1}^{N(u)} H_i > u \right) \leq P \left( \sum_{i=1}^{\lfloor KT\lfloor u \lfloor T(u) \rfloor \rfloor} H_i > u \right) + P \left( N(u) \geq \lfloor KT(u) \rfloor \right) .
$$
Due to [5 Prop. 1],
\[
\lim_{u \to \infty} \frac{1}{\sqrt{u}} \log \mathbb{P} \left( \sum_{i=1}^{[KT(u)]} H_i > u \right) \leq -T^* J \left( \frac{1}{(T^*)^2} \right).
\]
Also, \( \mathbb{P} (N(u) \geq [KT(u)]) \) decays exponentially. This implies the upper bound.

We now consider case (ii), that is, \( T(u) = \beta u \) for \( \beta < 1/\mathbb{E}Q \). Replace the process \( Q \) by a process \( \bar{Q}^{(\varepsilon)} \), as follows; fix an \( \varepsilon > 0 \). Start in stationarity, and let \( \bar{Q}^{(\varepsilon)} \) equal \( Q \) until the buffer idles, say at time \( \tau \). Then let \( \bar{Q}^{(\varepsilon)} \) be set to \( \varepsilon \), and let it follow the workload process from then on with the same driving Lévy process as \( Q \); again, as soon as it hits zero, it is set to \( \varepsilon \), etc. Clearly, \( \bar{Q}^{(\varepsilon)} \geq Q \), but, in addition, \( 0 \leq \bar{Q}^{(\varepsilon)} - Q \leq \varepsilon \) for all \( r \), which also entails that \( 0 \leq \mathbb{E}Q^{(\varepsilon)} - \mathbb{E}Q \leq \varepsilon \).

Now the proof of case (i) can be mimicked, the only term to be still taken care of is, with \( K := (1 + \delta)/\mathbb{E}\tau^{(\varepsilon)} \),
\[
\mathbb{P} \left( H + \sum_{i=1}^{[KT(u)]} H^{(\varepsilon)}_i > u - \beta u(1 + \delta) \frac{\mathbb{E}H^{(\varepsilon)}}{\mathbb{E}\tau^{(\varepsilon)}} \right).
\]
Choosing \( \delta, \varepsilon \) sufficiently small, the requirement \( \beta < 1/\mathbb{E}Q \) entails that
\[
1 - \beta(1 + \delta) \frac{\mathbb{E}H^{(\varepsilon)}}{\mathbb{E}\tau^{(\varepsilon)}} > 0,
\]
realizing that \( \mathbb{E}H^{(\varepsilon)}/\mathbb{E}\tau^{(\varepsilon)} \to \mathbb{E}Q \). Then the upper bound follows as in the proof of Theorem 1 in [5].

6. Final Remarks

We note that we can dispense from Assumption 3 as follows. Suppose that \( X \) only satisfies Assumptions 1 and 2. Let us consider the standard decomposition
\[
X_t = \mu t + \sigma B_t + Y_t,
\]
where \((Y_t)_{t \in \mathbb{R}}\) is a pure jump process independent of the Brownian motion \((B_t)_{t \in \mathbb{R}}\), \( \mu \) is a drift parameter and \( \sigma \geq 0 \) is a volatility parameter. Note that by truncating the negative jumps of \((Y_t)_{t \in \mathbb{R}}\) (i.e., keeping only those whose size is smaller \( c \)) we obtain a process \((Q^c_t)_{t \in \mathbb{R}}\) such that \( Q^c_t \geq Q_t \). The whole large deviations analysis can be performed under \((Q^c_t)_{t \in \mathbb{R}}\), which now satisfies Assumptions 1 and 2. Then finally one can let \( c \to \infty \).

It is natural to consider, particularly in the regime studied in Section 5, the asymptotics assuming that \( Q_0 = 0 \). It is easily shown that under Assumptions 1 and 2, the asymptotic
results derived in Section 5 remain unchanged. To see this note that the analysis obtained provides a correct upper bound for the asymptotics (by monotonicity, bearing in mind that starting in stationarity dominates the zero initial condition). To show the lower bound one simply isolates a path that falls within the event under consideration. Note that since \( a^* = 0 \), as discussed in Remark 1, one can consider a path for which \( Q_0 \leq \epsilon \sqrt{u} \) and the area reaches level \( \beta u \) in the first busy period. Since \( \epsilon \) is arbitrary one thus obtain that the large deviations asymptotics hold also for the case \( Q_0 = 0 \) as claimed.

7. APPENDIX: PROOF OF LEMMA 2

We start by quoting the following result due to [7] which is immediately applicable under Assumptions 2 and 3.

Lemma 4. If there exists \( \delta > 0 \) such that \( \mathbb{E} \exp (\delta |X_1|) < \infty \), then the process \( X_u/u \) satisfies a large deviations principle under the topology of convergence on compact sets.

We need to extend the previous result to accommodate the topology that we use, namely that generated by the norm \( \|\cdot\| = \sup_{0 \leq t < \infty} |f(t)| / (1 + t) \). Let us define \( L_1[0,\infty) \) to be the space of continuous functions endowed with the topology generated by the norm \( \|\cdot\| \).

A sequence of probability measures \( (\mathbb{P}_n)_{n \in \mathbb{N}} \) is exponentially tight if for every \( \lambda > 0 \) there exists a compact set \( K_\lambda \) such that

\[
\lim_{n \to \infty} \mathbb{P}_n(K_\lambda) \leq -\lambda.
\]

Exponential tightness is a crucial concept in order to lift the upper bound in the sample-path large deviations principle from compact sets to all closed sets. The next result characterizes exponential tightness in \( L_1[0,\infty) \).

Lemma 5. Consider a sequence of probability measures \( (\mathbb{P}_n)_{n \in \mathbb{N}} \) on \( L_1[0,\infty) \) (which are such that \( \mathbb{P}_n(x : x(0) = 0) = 1 \)) and acting on the Borel sigma-field corresponding to the topology generated by the norm \( \|\cdot\| \). Then, \( (\mathbb{P}_n)_{n \in \mathbb{N}} \) is exponentially tight if and only if: a) \( (\mathbb{P}_n)_{n \in \mathbb{N}} \) is exponentially tight under the relative topology generated by uniform convergence on compact sets (called Stone’s topology), and b) for each \( \delta > 0 \),

\[
\lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}_n \left( x : \sup_{t \geq t_0} \frac{|x(t)|}{t} > \delta \right) \to -\infty \text{ as } t_0 \to \infty.
\]

Proof. [10] Lemma 3.3] establishes that relatively compact sets in \( (L_1[0,\infty), \|\cdot\|_1) \) are those sets \( B \) with compact closure under the relative Stone topology, and satisfying

\[
\lim_{t \to \infty} \sup_{x \in B} \frac{|x(t)|}{t} = 0.
\]

Also, recall that (if \( x(0) = 0 \) a.s. with respect to each \( \mathbb{P}_n \)) for exponential tightness under Stone’s topology, it is necessary and sufficient (see [9] p. 30)) that, for each \( \epsilon, T > 0 \),

\[
\lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}_n \left( x : \omega(x, \delta, T) > \epsilon \right) \to -\infty \text{ as } \delta \downarrow 0,
\]
where \( \omega(x, \delta, T) \) is the modulus of continuity of \( x \), on the interval \([0, T]\), evaluated at \( \delta \).

We now show that if conditions (12) and (13) are satisfied, then the sequence \( \{P_n\}_{n \in \mathbb{N}} \) is exponentially tight. Pick \( \lambda > 0 \), choose \( \delta_k \) so that

\[
P_n \left( x : \omega(x, \delta_k, T) > 1/k \right) \leq e^{-n\lambda/2^{k+1}},
\]

and let \( B_k = \{ x : \omega(x, \delta_k, T) \leq 1/k \} \). Also, pick \( t_k \) so that

\[
P_n \left( x : \sup_{t \geq t_k} \frac{|x(t)|}{t} > 1/k \right) \leq e^{-n\lambda/2^{k+1}},
\]

and let \( C_k = \{ x : \sup_{t > t_k} |x(t)|/t \leq 1/k \} \). Consider the closure, \( \bar{A}_\lambda \), of \( A_\lambda := \cap_k (B_k \cap C_k) \).

Note that

\[
1 - P(\bar{A}_\lambda) \leq 1 - P(A_\lambda) = P(\cup_k (B_k^c \cap C_k^c)) \leq e^{-n\lambda}
\]

We now claim that \( A_\lambda \) is relatively compact (i.e., that \( \bar{A}_\lambda \) is compact). To see this, choose \( \varepsilon > 0 \) and let \( k_0 > 1/\varepsilon \). Then, for all \( \delta < \delta_{k_0} \) we have that

\[
\sup_{x \in A} \omega(x, \delta, T) < \varepsilon.
\]

Similarly, for every \( T > t_{k_0} \) we have that

\[
\varepsilon > \sup_{x \in A} \sup_{t > T} \frac{|x(t)|}{t},
\]

which implies that

\[
\lim_{t \to \infty} \sup_{x \in A} \frac{|x(t)|}{t} \leq \varepsilon
\]

for all \( \varepsilon > 0 \). Thus, by virtue of the Arzelà-Ascoli theorem (see [4, p. 81] and [10, Lemma 3.3]), we conclude the argument for sufficiency. The necessity part is easier and it follows just as in [9, p. 30]. Therefore, it is omitted. \( \square \)

We now are ready to provide the proof of Lemma 2.

**Proof of Lemma 2** Using a discrete polygonal approximation similar to the one used in the proof of [7, Thm. 1.2], we can introduce a continuous exponential approximation to \( X_{u}/u \). So, we might assume that the underlying space is \( L_1[0, \infty) \). In view of Lemma 4, in order to lift the large deviations result to the topology generated by \( \|\cdot\| \), it suffices to verify, according to [8, Corollary 4.2.6], that \( X_{u}/u \) is exponentially tight. Because the rate function \( I(\cdot) \) is known to have compact level sets, Lemma 4 implies exponential tightness of \( X_{u}/u \) under the uniform topology on compact sets. Because of Lemma 5, it then suffices to verify that, for any \( \delta > 0 \),

\[
\lim \sup_{u \to \infty} \log \mathbb{P} \left( \sup_{t > t_0} \left| \frac{X_{ut}/u - t \mathbb{E}X_1}{t} \right| \geq \delta \right) \to -\infty \quad \text{as} \quad t_0 \to \infty.
\]

Note that for any \( 0 < a < b < \infty \), the mapping \( x \mapsto \sup_{t \in [a,b]} |x(t)|/t \) is continuous under the topology generated by \( \|\cdot\| \). This implies that the family \( V_u = \sup_{t \in [a,b]} |X_{tu}/u - t \mathbb{E}X_1| \)
satisfies an LDP with rate function $H(\cdot)$, say. Hence, we can write
\[
\mathbb{P}\left(\sup_{t > t_0} \left| \frac{X_{tu}/u - t\mathbb{E}X_1}{t} \right| \geq \delta \right) \leq \sum_{k=1}^{\infty} \mathbb{P}\left( \sup_{t \in [t_0, t_0+k]} \left| \frac{X_{tu}/u - k\mathbb{E}X_1}{k} \right| \geq \delta \right) = \sum_{k=1}^{\infty} \exp\left(-H(\delta) + o_{kt_0u}(1)\right),
\]
where the subindex in $o_{kt_0u}(1)$ has been used just to emphasize that $o_{kt_0u}(1) \to 0$ as $kt_0u \to \infty$. So we can choose $k_0$ big enough so that for every $k > k_0$ we have $H(\delta) + o_{kt_0u}(1) > H(\delta)/2 > 0$. From these estimates the statement of the lemma follows. \qed

Acknowledgement: We thank the Editors and the Referees for their careful reviews and their useful recommendations. Blanchet acknowledges support from the NSF through the grants CMMI-0846816 and 1069064.

REFERENCES