J. Appl. Prob. **52,** 68–81 (2015)
Printed in England

LOGARITHMIC ASYMPTOTICS FOR MULTIDIMENSIONAL EXTREMES UNDER NONLINEAR SCALINGS

K. M. KOSIŃSKI,* *University of Amsterdam, and Eindhoven University of Technology*
M. MANDJES,** *University of Amsterdam, Eindhoven University of Technology, and CWI Amsterdam*

Abstract

Let $\boldsymbol{W} = \{\boldsymbol{W}_n : n \in \mathbb{N}\}$ be a sequence of random vectors in $\mathbb{R}^d$, $d \geq 1$. In this paper we consider the logarithmic asymptotics of the extremes of $\boldsymbol{W}$, that is, for any vector $\boldsymbol{q} > \boldsymbol{0}$ in $\mathbb{R}^d$, we find that $\log \mathbb{P}(\exists n \in \mathbb{N} \colon \boldsymbol{W}_n > u\boldsymbol{q})$ as $u \to \infty$. We follow the approach of the restricted large deviation principle introduced in Duffy (2003). That is, we assume that, for every $\boldsymbol{q} \geq \boldsymbol{0}$, and some scalings $\{a_n\}$, $\{v_n\}$, $(1/v_n) \log \mathbb{P}(\boldsymbol{W}_n/a_n \geq u\boldsymbol{q})$ has a, continuous in $\boldsymbol{q}$, limit $J_{\boldsymbol{W}}(\boldsymbol{q})$. We allow the scalings $\{a_n\}$ and $\{v_n\}$ to be regularly varying with a positive index. This approach is general enough to incorporate sequences $\boldsymbol{W}$, such that the probability law of $\boldsymbol{W}_n/a_n$ satisfies the large deviation principle with continuous, not necessarily convex, rate functions. The equations for these asymptotics are in agreement with the literature.

Keywords: Extrema of stochastic process; large deviation theory

2010 Mathematics Subject Classification: Primary 60F10
Secondary 60G70

1. Introduction

Let $W = \{W_n : n \in \mathbb{N}\}$ be a sequence of random variables taking values in $\mathbb{R}$. Define $Q = \sup_{n \geq 0} W_n$. The random variable Q has been extensively studied: if W is time-reversible, then Q has the same distribution as the steady-state workload distribution in a queue with free process W (see, e.g. Reich (1958)); Q has also various relations with finance and insurance risk. It is in general difficult to determine the distribution of Q. One could therefore settle for the less ambitious goal of identifying the corresponding tail asymptotics, that is, finding a function f, such that $\mathbb{P}(Q > u) \sim f(u)$ as $u \to \infty$ (i.e. the ratio of the two tends to 1 as $u \to \infty$). This, however, requires us to impose a quite restrictive structure on W, even in the Gaussian setting. Therefore, one usually resorts to determining the *logarithmic* asymptotics of the (right) tail of the distribution of Q. It has been observed that, in great generality,

$$\log \mathbb{P}(Q > u) = \log \mathbb{P}(\exists n \in \mathbb{N} \colon W_n > u) \sim \log \sup_{n \geq 0} \mathbb{P}(W_n > u) \quad \text{as } u \to \infty. \tag{1}$$

The heuristic behind this claim is the *principle of the largest term*: rare events occur in the most likely way. That is to say, if W is unlikely to ever reach level u, then conditional on W in fact reaching u, with overwhelming probability this happens near to the most likely epoch.

Received 1 November 2013; revision received 17 February 2014.
* Postal address: Instytut Matematyczny, University of Wrocław, pl. Grunwaldzki 2/4, 50–384 Wrocław, Poland. Email address: kosinski@math.uni.wroc.pl
** Postal address: Korteweg-de Vries Institute for Mathematics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, The Netherlands.

Using (1), the tail behavior of Q can be derived from the large deviation behavior of W; this was originally proposed by Kesidis *et al.* (1993) and later made rigorous by Glynn and Whitt (1994). More formally, let Λ be the limiting *cumulant generating function* (CGF) of W, that is, $\Lambda(\theta) = \lim_{n\to\infty} n^{-1} \log \mathbb{E} \exp(\theta W_n)$ when the limit exists, and let Λ satisfy the assumptions of the Gärtner–Ellis theorem. Then, by its virtue, the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, where μ_n is the law of W_n/n, satisfies the large deviation principle (LDP) with rate function Λ^*, the Fenchel–Legendre transform of Λ (also known as the convex conjugate of Λ); see Dembo and Zeitouni (1998) for the background on large deviations theory. Under these assumptions, Glynn and Whitt (1994) assert that

$$\lim_{u\to\infty} \frac{1}{u} \log \mathbb{P}(\exists n \in \mathbb{N} \colon W_n > u) = -\sup\{\theta \colon \Lambda(\theta) \le 0\}. \tag{2}$$

Owing to its generality, this result is useful in a broad range of applications.

The result in Glynn and Whitt (1994) has been extended in a notable paper by Duffield and O'Connell (1995) (see also Lelarge (2008)). The authors consider the logarithmic asymptotics of the tail of Q, by imposing assumptions on random variables W_n/a_n, where $\{a_n : n \in \mathbb{N}\}$ is some (not necessarily linear) scaling. It is assumed that the *scaled* limiting CGF of W, defined as $\Lambda(\theta) = \lim_{n\to\infty} v_n^{-1} \log \mathbb{E} \exp(\theta v_n W_n/a_n)$ for some sequence $\{v_n : n \in \mathbb{N}\}$, exists as an extended real number. The considered class of admissible scaling sequences is quite broad. It includes the case when $a \in \mathcal{RV}(A)$, $v \in \mathcal{RV}(V)$ are two regularly varying sequences with indices $A, V > 0$; this class is broad enough for most of the applications. Considering nonlinear scalings allows us to incorporate, for instance, long/short range dependent sequences corresponding to, for example, fractional Brownian motion. Similarly to Glynn and Whitt (1994), it is assumed that Λ meets the assumptions of the Gärtner–Ellis theorem. Consequently, the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, where μ_n is the law of W_n/a_n, satisfies the LDP with *speed* v_n and rate function Λ^*. The main result from Duffield and O'Connell (1995) states that, under some additional assumptions,

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N} \colon W_n > u) = -\inf_{c>0} c^{-V/A} \Lambda^*(c), \tag{3}$$

where $h = v \circ a^{-1}$ and a^{-1} denotes the right inverse of a. It can be verified that (2) follows from (3) in case of the linear scaling: $a_n = v_n = n$.

Both Duffield and O'Connell (1995) and Glynn and Whitt (1994) impose additional conditions on W only to infer that the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, whether μ_n is the law of W_n/n or W_n/a_n, satisfies the LDP with some well-behaved rate function. By exploiting the Gärtner–Ellis theorem, the considered class of possible rate functions is limited to convex functions, whereas, in general, it could be any lower semicontinuous function. Alternatively, therefore, one could assume that the LDP holds without any knowledge of how it was inferred. Duffy *et al.* (2003) suggested a variant of this approach that allows us to consider scaled sequences W with nonconvex rate functions. More formally, the authors proposed the *restricted* large deviation principle (RLDP) instead of the classical LDP. That is, they required that the limit

$$\lim_{n\to\infty} \frac{1}{v_n} \log \mathbb{P}\bigg(\frac{W_n}{a_n} > c\bigg) = -J_W(c) \tag{4}$$

exists for every $c \ge 0$ and the function J_W is continuous on the interior of the set upon which it is finite. Thus, if the LDP holds with a continuous (where finite) rate function I_W (and speed v_n), then the RLDP holds with $J_W(c) = \inf_{x \ge c} I_W(x)$. Nevertheless, it is not required that I_W has been inferred from the Gärtner–Ellis theorem, nor, in principle, that the LDP holds at all.

https://doi.org/10.1239/jap/1429282607 Published online by Cambridge University Press

Observe that J_W does not give any information about the behavior of W for negative values so that in general the LDP is not even a prerequisite for the RLDP to hold. The main result of Duffy *et al.* (2003), under some additional assumptions, reads

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon\ W_n > u) = -\inf_{c>0} c^{-V/A} J_W(c). \tag{5}$$

It can be verified that if $J_W(c) = \inf_{x\geq c} I_W(x)$ and I_W is convex, then (5) reduces to (3) in the special case when $I_W = \Lambda^*$.

Note that (4) is a statement about the limiting behavior of $\mathbb{P}(W_n/a_n > c)$ with n growing large. This condition does not extract any information about W for specific values of n, in particular the initial values of W. The distribution of Q however, and, hence, the asymptotics as well, does involve the whole sequence W. It is therefore possible that the asymptotics of a *single* W_n could dominate those of Q. For instance, one can greatly alter Q by simply substituting W_0 with a properly chosen heavy-tailed random variable $\tilde{W}_0$. To exclude such a scenario, Duffy *et al.* (2003) introduced an additional, novel assumption referred to as the *uniform individual decay rate* hypothesis; see Section 2.2. Roughly speaking it prevents the sequence W from having an 'unusual' behavior for a single W_n. In fact, Duffy *et al.* (2003, Section 4) points out that this issue was overlooked by Duffield and O'Connell (1995). That is, in order for the results from Duffield and O'Connell (1995) to hold, one actually needs to impose further conditions. In the light of the generality of the result by Duffy *et al.* (2003), their paper can be treated as the most up to date treatment of the subject of logarithmic asymptotics for the supremum of a stochastic sequence.

In this paper we generalize and extend the result from Duffy *et al.* (2003) in multiple ways. Firstly, in Theorem 4 we show that, under the assumptions of Duffy *et al.* (2003), the sequence of probability measures $\{\mu_u^W : u \in \mathbb{R}_+\}$, where $\mu_u^W(A) = \mathbb{P}(\exists n \in \mathbb{N}\colon\ W_n \in uA)$, satisfies the LDP with speed $h(u)$ and rate function $\tilde{I}_W$, such that $\tilde{I}_W(x) = x^{V/A}$ for $x \geq 0$, and $\tilde{I}_W(x) = \infty$ for $x < 0$. In particular, for any $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon\ W_n \in uA) = -\inf_{x\in A} x^{V/A} \inf_{c>0} c^{-V/A} J_W(c). \tag{6}$$

We can see that (6) extends (5) by setting $A = (1, \infty)$. Theorem 4 is presented in Section 3.1.

Furthermore, in Section 3.2 we allow the sequence W to take values in $\mathbb{R}^d$ for any $d \geq 1$, rather than just $\mathbb{R}$. As it turns out, this multidimensional setting imposes substantial additional challenges, as compared to the single-dimensional setting. Regarding notation, to explicitly distinguish the multidimensional case from the one-dimensional counterpart we will make use of the usual boldface fonts. That is, we write $\boldsymbol{x}$ for the vector $\boldsymbol{x} = (x_1, \dots, x_d)$, where the dimension d should be clear from the context. All vector relations should be understood coordinatewise; for instance, we write $\boldsymbol{v} \geq \boldsymbol{w}$ to mean $v_i \geq w_i$ for all $i = 1, \dots, d$. With this notation, we consider a sequence $\boldsymbol{W} = \{\boldsymbol{W}_n : n \in \mathbb{N}\}$ of random vectors in $\mathbb{R}^d$. The sequence $\boldsymbol{W}$ is assumed to satisfy multidimensional analogues of the assumptions from Duffy *et al.* (2003); see Section 2.2. In particular, it is assumed that the (multidimensional) RLDP holds, that is, the limit

$$\lim_{n\to\infty} \frac{1}{v_n} \log \mathbb{P}\left(\frac{\boldsymbol{W}_n}{a_n} > \boldsymbol{q}\right) = -J_{\boldsymbol{W}}(\boldsymbol{q})$$

exists for any $\boldsymbol{q} \geq \boldsymbol{0}$. Our second contribution, Theorem 5, states that, for any vector $\boldsymbol{q} > \boldsymbol{0}$,

$$\lim_{n\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon\ \boldsymbol{W}_n > u\boldsymbol{q}) = -\inf_{c>0} c^{-V/A} J_{\boldsymbol{W}}(c\boldsymbol{q}). \tag{7}$$

Obviously (7) is a generalization of (5) in the multidimensional sense.

Another significant contribution is that we also discuss various relations and connections with the existing literature. In Section 4.1 we present the relation between the RLDP approach, as undertaken here and in Duffy *et al.* (2003), and the approach via the CGFs, as undertaken in Duffield and O'Connell (1995) and Glynn and Whitt (1994). In Section 4.2 we discuss the various results of Collamore (1996), who considered a sequence of random vectors $\{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ in $\mathbb{R}^d$, such that the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, where μ_n corresponds to the law of $\boldsymbol{Y}_n/n$, satisfies the LDP with a convex rate function. Collamore proved various LDP-like statements for the sequence of probability measures $\{\mu_u^{\boldsymbol{Y},N} : u \in \mathbb{R}_+\}$, where $\mu_u^{\boldsymbol{Y},N}(A) = \mathbb{P}(\exists n \geq N : \boldsymbol{Y}_n \in uA)$. These results, not referred to in Duffy *et al.* (2003), coincide with (5)–(7) in the case of $N = 1$, linear scaling and convex rate function. We provide a discussion of these results also in Section 4.2.

We conclude our paper with Section 5, where we present an extension of Theorem 5 from sequences $\{\boldsymbol{W}_n : n \in \mathbb{N}\}$ to stochastic processes $\{\boldsymbol{W}_t : t \in \mathbb{R}_+\}$. Furthermore, we apply our results to two examples. In the first example we treat heavy-tailed processes which exhibit nonconvex rate functions, an example that was not covered by results that were known so far. In the second example we compare our results with Dębicki *et al.* (2010), which addresses a similar problem for the case of $\boldsymbol{W}$ being Gaussian.

2. Preliminaries

In this paper we use the following notation. For a function $f : \mathbb{R}^d \to \mathbb{R}$ we denote its domain by $\mathcal{D}_f = \{\boldsymbol{x} : f(\boldsymbol{x}) < \infty\}$. As already introduced in Section 1, we shall work with the following two functions $a, v : \mathbb{R}_+ \to \mathbb{R}_+$, which throughout the whole paper are assumed to be regularly varying functions at infinity with positive indices A and V, respectively; we write $a \in \mathcal{RV}(A)$ and $v \in \mathcal{RV}(V)$. It is well known that, for any regularly varying function $f \in \mathcal{RV}(F)$ with a positive index F, it is possible to construct a strictly increasing and continuous function f' such that

$$\lim_{t\to\infty} \frac{f'(ct)}{f'(t)} = \lim_{t\to\infty} \frac{f(ct)}{f(t)} = \lim_{t\to\infty} \frac{f'(ct)}{f(t)} = c^F.$$

Therefore, without loss of generality, we assume that both a and v are continuous and strictly increasing. We shall also speak about regularly varying sequences a_n and v_n defined by $a_n = a(n)$ and $v_n = v(n)$. Let a^{-1} denote the inverse of a. Define a new function $h : \mathbb{R}_+ \to \mathbb{R}_+$ by $h = v \circ a^{-1}$. The function h belongs to the class $\mathcal{RV}(V/A)$. For details on regular variation, see, e.g. Bingham *et al.* (1987).

For any subset A of $\mathbb{R}^d$ we denote its cone by $\operatorname{cone}(A) = \{\lambda\boldsymbol{x} : \boldsymbol{x} \in A,\ \lambda \geq 0\}$, its closure by $\bar{A}$, and interior by A°. For any convex function f on $\mathbb{R}^d$, we define its Fenchel–Legendre transform f^* as $f^*(\boldsymbol{x}) = \sup_{\boldsymbol{\alpha}\in\mathbb{R}^d}(\langle\boldsymbol{\alpha}, \boldsymbol{x}\rangle - f(\boldsymbol{\alpha}))$.

2.1. Large deviations theory

We follow the definitions and setup as used in Dembo and Zeitouni (1998). All probability measures in this paper are assumed to be Borel measures. The function $I : \mathbb{R}^d \to [0, \infty]$ is called a *rate function* if I is lower semicontinuous and $I \not\equiv \infty$. We say that I is a *good* rate function if, in addition, the level sets $\mathcal{L}_a I = \{\boldsymbol{x} : I(\boldsymbol{x}) \leq a\}$ are compact for each $a \geq 0$ (which in fact is equivalent to the level sets being bounded as a function f is lower semicontinuous if and only if its level sets are closed). A sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$ on $\mathbb{R}^d$ is said to satisfy the (LDP) with *rate function* I if, for all $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$-\inf_{x\in\Gamma^\circ} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(\Gamma) \leq \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(\Gamma) \leq -\inf_{x\in\bar{\Gamma}} I(x). \tag{8}$$

https://doi.org/10.1239/jap/1429282607 Published online by Cambridge University Press

Consider the empirical mean $\boldsymbol{Z}_n = (1/n)\sum_{i=1}^n \boldsymbol{X}_i$ of a sequence of independent and identically distributed (i.i.d.) random vectors $\{\boldsymbol{X}_n : n \in \mathbb{N}\}$ in $\mathbb{R}^d$. Define μ_n as the law of $\boldsymbol{Z}_n$ and let $\Lambda_X = \log \mathbb{E}\mathrm{e}^{\langle \boldsymbol{\alpha}, \boldsymbol{X}_1 \rangle}$ be the (CGF) associated to the law of $\boldsymbol{X}_1$. In this case the classical theorem of Cramér applies.

Theorem 1. (Cramér's theorem.) *Assume that* $0 \in \mathcal{D}^\circ_{\Lambda_X}$, *then* $\{\mu_n : n \in \mathbb{N}\}$ *satisfies the LDP on* $\mathbb{R}^d$ *with good rate function* Λ^*_X.

Now consider a general sequence of random vectors $\{\boldsymbol{Z}_n : n \in \mathbb{N}\}$ in $\mathbb{R}^d$, and let μ_n again denote the law of $\boldsymbol{Z}_n$. The CGF associated with the law μ_n is defined as $\Lambda_n(\boldsymbol{\alpha}) = \log \mathbb{E}\mathrm{e}^{\langle \boldsymbol{\alpha}, \boldsymbol{Z}_n \rangle}$. Let $\Lambda(\boldsymbol{\alpha}) = \limsup_{n\to\infty}(1/n)\Lambda_n(n\boldsymbol{\alpha})$ be the limiting CGF. With this notation, the following well-known theorem holds.

Theorem 2. (Gärtner–Ellis theorem.) *Assume that* $0 \in \mathcal{D}^\circ_\Lambda$ *and* Λ *is an essentially smooth, lower semicontinuous function. Then* $\{\mu_n : n \in \mathbb{N}\}$ *satisfies the LDP on* $\mathbb{R}^d$ *with good rate function* Λ^*.

Note that if $\boldsymbol{Z}_n = (1/n)\sum_{i=1}^n \boldsymbol{X}_i$ as in the setting of Cramér's theorem, then $\Lambda = \Lambda_X$ and further regularity conditions are not required.

We can consider LDPs with the so-called *speed* $\{s_n : n \in \mathbb{N}\}$ when $1/n$ in (8) is replaced by $1/s_n \to 0$. The Gärtner–Ellis theorem remains valid if $\Lambda(\boldsymbol{\alpha}) = \limsup_{n\to\infty}(1/s_n)\Lambda_n(s_n\boldsymbol{\alpha})$, the scaled limiting CGF, satisfies the assumptions. All results of this subsection carry through to continuous parameter families $\{\mu_u : u \in \mathbb{R}_+\}$.

2.2. Main assumptions

In this paper we consider sequences $\boldsymbol{W} = \{\boldsymbol{W}_n : n \in \mathbb{N}\}$ of random vectors in $\mathbb{R}^d$, $d \geq 1$, satisfying the following assumptions.

Assumption 1. (Restricted LDP hypothesis.) *There exists a function* $J_{\boldsymbol{W}} : \mathbb{R}^d_+ \to [0, \infty]$ *such that, for every* $\boldsymbol{q} \geq \boldsymbol{0}$,

$$\lim_{n\to\infty} \frac{1}{v_n} \log \mathbb{P}\left(\frac{\boldsymbol{W}_n}{a_n} > \boldsymbol{q}\right) = -J_{\boldsymbol{W}}(\boldsymbol{q}). \tag{9}$$

Remark 1. If the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, where μ_n denotes the law of $\boldsymbol{W}_n/a_n$, satisfies the LDP with speed v_n and rate function $I_{\boldsymbol{W}}$, which is continuous where it is finite, then it also satisfies the RLDP hypothesis with $J_{\boldsymbol{W}}(\boldsymbol{q}) = \inf_{\boldsymbol{x}\geq\boldsymbol{q}} I_{\boldsymbol{W}}(\boldsymbol{x})$ (hence the name of the hypothesis). If I is not continuous then it is easy to construct an example in which the restricted LDP does not hold. The opposite implication is also not true in general, that is, the restricted LDP hypothesis does not imply the LDP: property (9) does not provide any information about the negative values of $\boldsymbol{W}_n/a_n$.

Assumption 2. (Stability and continuity hypothesis.) *It holds that* $J_{\boldsymbol{W}}(\boldsymbol{0}) > 0$ *and there exists* $\boldsymbol{y} > \boldsymbol{0}$ *such that* $J_{\boldsymbol{W}}(\boldsymbol{y}) < \infty$. *Furthermore,* $J_{\boldsymbol{W}}$ *is assumed to be continuous on* $\mathcal{D}^\circ_{J_{\boldsymbol{W}}}$.

In the queueing context, $J_{\boldsymbol{W}}(\boldsymbol{0}) > 0$ is the usual stability condition. If $J_{\boldsymbol{W}}(\boldsymbol{x}) = \infty$ for all $\boldsymbol{x} \in \mathbb{R}^d_+ \setminus \{\boldsymbol{0}\}$ then $\mathbb{P}(\exists n \in \mathbb{N} : \boldsymbol{W}_n > u\boldsymbol{q})$ will have superexponential decay.

The restricted LDP hypothesis refers to the limiting behavior of $\log \mathbb{P}(\boldsymbol{W}_n > ua_n\boldsymbol{q})$ for large n, not values for specific n. The asymptotic of a single $\boldsymbol{W}_n$ could dominate those of $\mathbb{P}(\exists n \in \mathbb{N} : \boldsymbol{W}_n > u\boldsymbol{q})$. The condition below excludes this possibility.

Condition 1. (Uniform individual decay rate hypothesis.) *For a fixed vector* $\boldsymbol{q} > \boldsymbol{0}$*, there exist constants* $F = F(\boldsymbol{q}) > V/A$ *and* $K = K(\boldsymbol{q}) > 0$ *so that, for all* n *and all* $c > K$*,*

$$\frac{1}{v_n} \log \mathbb{P}\bigg(\frac{\boldsymbol{W}_n}{a_n} > c\boldsymbol{q} \bigg) \le -c^F.$$

Remark 2. The restricted LDP hypothesis, the stability and continuity hypothesis, and the uniform individual decay rate hypothesis were originally introduced in Duffy *et al.* (2003) in the one-dimensional case. That is, if we set $d = 1$ then all the above hypotheses reduce to those of Duffy *et al.* (2003). The hypotheses presented above can therefore be seen as natural extensions of the hypotheses from Duffy *et al.* (2003) to the multidimensional setting.

Theorem 3. (One-dimensional case Duffy *et al.* (2003, Theorem 2.2).) *If the sequence* $W = \{W_n : n \in \mathbb{N}\}$ *of random variables satisfies all the hypotheses of this subsection then*

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N} \colon W_n > u) = -\inf_{c>0} c^{-V/A} J_W(c). \tag{10}$$

3. Two extensions

In this section we present two generalizations of Theorem 3. Firstly, we consider families of measures $\mu_u^W(A) = \mathbb{P}(\exists n \in \mathbb{N} \colon W_n \in uA)$ for a general set A; we can easily see that Theorem 3 considers the case of $A = (1, \infty)$. Secondly, we consider the situation when the sequence W takes values in $\mathbb{R}^d$, $d \ge 1$, rather than just $\mathbb{R}$.

3.1. Extension to the LDP

Define a lower semicontinuous function $\tilde{I}_W$ by

$$\tilde{I}_W(x) = \begin{cases} \infty & \text{for } x < 0, \\ k_{J_W} x^{V/A} & \text{for } x \ge 0, \end{cases}$$

where $k_{J_W} = \inf_{c>0} c^{-V/A} J_W(c)$ is the constant appearing on the right-hand side of (10). In this subsection we assume that the sequence of random variables $W = \{W_n : n \in \mathbb{N}\}$ is such that $W_0 = 0$. This assures that $Q = \sup_{n\ge 0} W_n$ is a nonnegative random variable. Let μ_u^W be a probability measure on $\mathbb{R}$ defined as $\mu_u^W(A) = \mathbb{P}(\exists n \in \mathbb{N} \colon W_n \in uA)$. The following theorem is the main result of this subsection.

Theorem 4. *Under the assumptions of Theorem 3, the family* $\{\mu_u^W : u \in \mathbb{R}_+\}$ *satisfies the LDP with speed function* h *and good rate function* $\tilde{I}_W$.

Proof. First note that, for any $k > 0$,

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mu_u^W((k, \infty)) = -\tilde{I}_W(k). \tag{11}$$

Indeed,

$$\frac{1}{h(u)} \log \mu_u^W((k, \infty)) = \frac{h(ku)}{h(u)} \frac{1}{h(ku)} \log \mu_{ku}^W((1, \infty)),$$

so that (11) follows from Theorem 3 combined with the fact that $h(ku)/h(u)$ tends to $k^{V/A}$.

https://doi.org/10.1239/jap/1429282607 Published online by Cambridge University Press

Let Γ be any set in $\mathcal{B}(\mathbb{R})$. Now, if for some y, $\inf_{x\in\bar{\Gamma}} \tilde{I}_W(x) = \tilde{I}_W(y)$ then, for any $\eta \in (0, y)$, the monotonicity of $\tilde{I}_W$ on $\mathbb{R}_+$ implies that

$$\mu_u^W(\Gamma) \le \mu_u^W([y, \infty)) \le \mu_u^W((y-\eta, \infty)).$$

Hence, by (11),

$$\limsup_{u\to\infty} \frac{1}{h(u)} \log \mu_u^W(\Gamma) \le -\tilde{I}_W(y-\eta).$$

This combined with the continuity of $\tilde{I}_W$ on $\mathbb{R}_+$ implies the upper bound:

$$\limsup_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon W_n \in u\Gamma) \le -\inf_{x\in\bar{\Gamma}} \tilde{I}_W(x).$$

If $\inf_{x\in\bar{\Gamma}} \tilde{I}_W(x) = 0$ or $\inf_{x\in\bar{\Gamma}} \tilde{I}_W(x) = \infty$, then the above bound holds trivially.

Now, if $\Gamma^\circ \cap \mathbb{R}_+ = \varnothing$ then the lower bound

$$\liminf_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon W_n \in u\Gamma) \ge -\inf_{x\in\Gamma^\circ} \tilde{I}_W(x)$$

holds immediately. Therefore, let $\Gamma^\circ \subset \mathbb{R}_+$ and $s \in \Gamma^\circ$, $\eta > 0$ be such that $(s-\eta, s+\eta] \subset \Gamma^\circ$. Obviously, $\mu_u^W(\Gamma) \ge \mu_u^W((s-\eta, s+\eta]) = \mu_u^W((s-\eta,\infty)) - \mu_u^W((s+\eta,\infty))$. From (11), for sufficiently large u, $\mu_u^W((s+\eta,\infty)) \le \mu_u^W((s-\eta,\infty))/2$. Hence,

$$\liminf_{u\to\infty} \frac{1}{h(u)} \log \mu_u^W(\Gamma) \ge \liminf_{u\to\infty} \frac{1}{h(u)} \log\bigg(\frac{\mu_u^W((s-\eta,\infty))}{2}\bigg) = -\tilde{I}_W(s-\eta) \ge -\tilde{I}_W(s).$$

Now the lower bound follows after optimization over $s \in \Gamma^\circ$.

Remark 3. The assumption that $W_0 = 0$ is natural and is fulfilled in many applications. If it does not hold, however, then the LDP from Theorem 4 remains true on $\mathcal{B}(\mathbb{R}_+)$. In both of these cases, the continuity of $\tilde{I}_W$ on $\mathbb{R}_+$ implies that for any set $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon W_n \in uA) = -\inf_{x\in A} \tilde{I}_W(x).$$

Hence, Theorem 4 generalizes Theorem 3, which can be seen by taking $A = (1, \infty)$.

3.2. Extension to the multidimensional case

In this subsection we generalize Theorem 3 to the multidimensional case. Theorem 3 itself serves as a building block in the proof of the following theorem.

Theorem 5. (Multidimensional case.) *If the sequence* $\boldsymbol{W} = \{\boldsymbol{W}_n : n \in \mathbb{N}\}$ *of random vectors in* $\mathbb{R}^d$, $d \ge 1$, *satisfies the hypotheses of Section 2.2, then, for any* $\boldsymbol{q} > \boldsymbol{0}$,

$$\lim_{u\to\infty} \frac{1}{h(u)} \log \mathbb{P}(\exists n \in \mathbb{N}\colon \boldsymbol{W}_n > u\boldsymbol{q}) = -\inf_{c>0} c^{-V/A} J_{\boldsymbol{W}}(c\boldsymbol{q}).$$

Proof. Note that $\mathbb{P}(\boldsymbol{W}_n > u\boldsymbol{q}) = \mathbb{P}(Z_n > u)$, where $Z = \{Z_n : n \in \mathbb{N}\}$ is a sequence of random variables such that $Z_n = \min_{i=1,\dots,d}(W_{n,i}/q_i)$. The sequence Z satisfies the assumptions of Theorem 3 with a function J_Z given by $J_Z(c) = J_{\boldsymbol{W}}(c\boldsymbol{q})$. Indeed, by the restricted LDP hypothesis, it follows that, for every $c \ge 0$,

$$\lim_{n\to\infty} \frac{1}{v_n} \log \mathbb{P}\bigg(\frac{Z_n}{a_n} > c\bigg) = \lim_{n\to\infty} \frac{1}{v_n} \log \mathbb{P}\bigg(\frac{\boldsymbol{W}_n}{a_n} > c\boldsymbol{q}\bigg) = -J_{\boldsymbol{W}}(c\boldsymbol{q}).$$

The stability and continuity hypothesis for J_Z easily follows from the stability and continuity hypothesis for $J_{\boldsymbol{W}}$. The same applies to the uniform individual decay rate hypothesis. Therefore, Theorem 3 yields

$$\begin{aligned}\lim_{u\to\infty}\frac{1}{h(u)}\log\mathbb{P}(\exists n\in\mathbb{N}\colon \boldsymbol{W}_n>u\boldsymbol{q})\\ &=\lim_{u\to\infty}\frac{1}{h(u)}\log\mathbb{P}(\exists n\in\mathbb{N}\colon Z_n>u)\\ &=-\inf_{c>0}c^{-V/A}J_Z(c)\\ &=-\inf_{c>0}c^{-V/A}J_{\boldsymbol{W}}(c\boldsymbol{q}).\end{aligned}$$

This completes the proof.

4. Connections with existing literature

We have already discussed the relations of our results to Duffy *et al.* (2003). In this section we shall discuss our findings in light of already existing results of Collamore (1996), Duffield and O'Connell (1995), and Glynn and Whitt (1994).

4.1. The CGF approach

The analyses in Duffield and O'Connell (1995) and Glynn and Whitt (1994) are based on the CGF. They both consider the $d=1$ case, but only Duffield and O'Connell (1995) allows for nonlinear scaling. In this subsection we present conditions under which the main assumptions of the present paper are fulfilled. Recall that,

$$\Lambda_n(\boldsymbol{\alpha})=\log\mathbb{E}\exp\biggl(\biggl\langle\boldsymbol{\alpha},\frac{\boldsymbol{W}_n}{a_n}\biggr\rangle\biggr)$$

is the CGF of the law of $\boldsymbol{W}_n/a_n$. Here it is assumed that Λ_n exists as a finite real number for all $\boldsymbol{\alpha}\in\mathbb{R}^d$ and all $n\in\mathbb{N}$. The hypotheses of our paper have simple expressions in terms of the CGF. The conditions we specify here for the CGF case, based on their one-dimensional analogues in Duffy *et al.* (2003), are intended for easy applicability rather than maximum generality. Under these assumptions, the large deviation rate function is convex. The CGF technique is not applicable to models which have nonconvex rate functions.

Assumption 3. (LDP hypothesis, CGF case.) *For each $\boldsymbol{\alpha}\in\mathbb{R}^d$, the scaled limiting CGF $\Lambda(\boldsymbol{\alpha})=\lim_{n\to\infty}(1/v_n)\Lambda_n(v_n\boldsymbol{\alpha})$ exists. Furthermore, Λ is assumed to be continuously differentiable.*

Under the above assumption, by the Gärtner–Ellis theorem, the sequence of probability measures $\{\mu_n\colon n\in\mathbb{N}\}$, where μ_n is the law of $\boldsymbol{W}_n/a_n$, satisfies the LDP with rate function $I_{\boldsymbol{W}}=\Lambda^*$ and speed v_n. This implies that $I_{\boldsymbol{W}}$ is convex and continuous on the set where it is finite and, therefore, the RLDP holds with $J_{\boldsymbol{W}}(\boldsymbol{q})=\inf_{\boldsymbol{x}\geq\boldsymbol{q}}\Lambda^*(\boldsymbol{x})$, which is also continuous on $\mathcal{D}^{\circ}_{J_{\boldsymbol{W}}}$. Hence, in order to assure that the stability and continuity hypothesis holds, we require the following conditions.

Condition 2. (Stability hypothesis, CGF case.) *There exists $\boldsymbol{\alpha}^\star>\boldsymbol{0}$ such that $\Lambda(\boldsymbol{\alpha}^\star)<0$.*

The above hypothesis implies that $J_{\boldsymbol{W}}(\boldsymbol{0})\geq-\Lambda(\boldsymbol{\alpha}^\star)>0$. Indeed,

$$J_{\boldsymbol{W}}(\boldsymbol{0})=\inf_{\boldsymbol{x}\geq\boldsymbol{0}}\sup_{\boldsymbol{\alpha}\in\mathbb{R}^d}(\langle\boldsymbol{\alpha},\boldsymbol{x}\rangle-\Lambda(\boldsymbol{\alpha}))\geq\sup_{\boldsymbol{\alpha}\in\mathbb{R}^d_+}\inf_{\boldsymbol{x}\geq\boldsymbol{0}}(\langle\boldsymbol{\alpha},\boldsymbol{x}\rangle-\Lambda(\boldsymbol{\alpha}))=-\inf_{\boldsymbol{\alpha}\in\mathbb{R}^d_+}\Lambda(\boldsymbol{\alpha})\geq-\Lambda(\boldsymbol{\alpha}^\star).$$

Condition 3. (Uniform individual decay rate hypothesis, CGF case.) *There exist constants* F' *and* M *such that* $F' > \max\{V/A, 1\}$ *and* $(1/v_n)\Lambda_n(v_n\boldsymbol{\alpha}) \leq M\|\boldsymbol{\alpha}\|^{F'/(F'-1)}$ *for all* $\boldsymbol{\alpha} > \mathbf{0}$ *and all* $n \in \mathbb{N}$.

Under this hypothesis, for each $F \in (1, F')$, there exists a constant $K_F = K_F(\boldsymbol{q})$ such that, for all $c > K_F$ and all $n \in \mathbb{N}$,

$$\frac{1}{v_n}\log\mathbb{P}\bigg(\frac{\boldsymbol{W}_n}{a_n} > c\boldsymbol{q}\bigg) \leq -c^F.$$

That is, the uniform individual decay rate hypothesis certainly holds. Indeed, an elementary consequence of Chernoff's inequality is

$$\log\mathbb{P}(\boldsymbol{W}_n > ca_n\boldsymbol{q}) \leq -v_n\bigg(c\langle\boldsymbol{\alpha}, \boldsymbol{q}\rangle - \frac{1}{v_n}\Lambda_n(v_n\boldsymbol{\alpha})\bigg)$$

for any $\boldsymbol{\alpha} > \mathbf{0}$. It then follows that

$$\log\mathbb{P}(\boldsymbol{W}_n > ca_n\boldsymbol{q}) \leq -v_n(c\langle\boldsymbol{\alpha}, \boldsymbol{q}\rangle - M\|\alpha\|^{F'/(F'-1)}).$$

Choosing $\boldsymbol{\alpha} = (c(F'-1)\|\boldsymbol{q}\|/(MF'))^{F'-1}\boldsymbol{q}$, we have

$$\log\mathbb{P}(\boldsymbol{W}_n > ca_n\boldsymbol{q}) \leq -v_n(c\|q\|^2)^{F'}(M^{1-F'}F'^{-F'}(F'-1)^{F'-1}). \tag{12}$$

Since M and F' are constants, for each $F \in (\max\{V/A, 1\}, F')$, there exists $K_F = K_F(\boldsymbol{q})$ such that, for all $c > K_F$, the right-hand side of (12) will be less than $-v_nc^F$.

4.2. The LDP with a convex rate function

The purpose of this subsection is to discuss the differences between the approach from Collamore (1996) and the one in our paper.

Collamore (1996) considered a sequence $\boldsymbol{Y} = \{\boldsymbol{Y}_n : n \in \mathbb{N}\}$ of random vectors in $\mathbb{R}^d$. The main assumption of his paper is that the sequence of probability measures $\{\mu_n : n \in \mathbb{N}\}$, where μ_n is the law of $\boldsymbol{Y}_n/n$, satisfies the LDP with a convex rate function $I_{\boldsymbol{Y}}$, such that $\mathcal{L}_0 I_{\boldsymbol{Y}} \neq \varnothing$, that is $I_{\boldsymbol{Y}}(\boldsymbol{x}) = 0$ for some $\boldsymbol{x} \in \mathbb{R}^d$. Furthermore, it is assumed that (using the notation from Collamore (1996)):

(H1'): $\sup_{n\geq N}(1/n)\log\mathbb{E}\exp\langle\boldsymbol{\alpha}, \boldsymbol{Y}_n\rangle < \infty$ for all $\boldsymbol{\alpha} \in \mathcal{L}_0 I_{\boldsymbol{Y}}^*$ and N greater than or equal to some N_0.

Remark 4. (H1') is a regularity condition on the sequence $\boldsymbol{Y}$. It is satisfied if $\boldsymbol{Y}$ is, for example, the nth partial sum of an i.i.d. sequence of random vectors satisfying the condition given in Cramér's theorem, or the weaker condition given by Ney and Robinson (1995). It also holds when $\boldsymbol{Y}$ is a Markov-additive process satisfying the uniform recurrence condition (6.2) of Ney and Nummelin (1987). In both these cases N_0 can be taken to be 1. Also, $\boldsymbol{Y}$ can be a general sequence satisfying the conditions of the Gärtner–Ellis theorem and (i) $\Lambda(\boldsymbol{\alpha})$ is finite in the neighborhood of each $\boldsymbol{\alpha} \in \mathcal{L}_0\Lambda$; (ii) the level sets of Λ are compact; recall that $\Lambda(\boldsymbol{\alpha}) = \limsup_{n\to\infty}(1/n)\log\mathbb{E}\exp\langle\boldsymbol{\alpha}, \boldsymbol{Y}_n\rangle$.

Theorem 6. (Collamore (1996, Theorem 2.2).) *Suppose* A *is a general set in* $\mathbb{R}^d$ *and (H1') and*

(H2): *for some* $\delta > 0$, $A \cap \operatorname{cone}(\mathcal{C}_\delta) = \varnothing$, *where* $\mathcal{C}_\delta = \{\boldsymbol{x} : \inf_{\boldsymbol{y}\in\mathcal{L}_0 I_{\boldsymbol{Y}}}\|\boldsymbol{x} - \boldsymbol{y}\| < \delta\}$;

https://doi.org/10.1239/jap/1429282607 Published online by Cambridge University Press

are satisfied. Then, for any $N \geq N_0$,

$$\liminf_{u\to\infty} \frac{1}{u} \log \mathbb{P}(\exists n \geq N \colon \boldsymbol{Y}_n \in uA) \geq -\inf_{\boldsymbol{x}\in A^\circ} \tilde{I}_{\boldsymbol{Y}}(\boldsymbol{x})$$

and

$$\limsup_{u\to\infty} \frac{1}{u} \log \mathbb{P}(\exists n \geq N \colon \boldsymbol{Y}_n \in uA) \leq -\inf_{\boldsymbol{x}\in \bar{A}} \tilde{I}_{\boldsymbol{Y}}(\boldsymbol{x}),$$

where $\tilde{I}_{\boldsymbol{Y}}(\boldsymbol{x}) = \sup_{\boldsymbol{\alpha}\in\mathcal{L}_0 I^*_{\boldsymbol{Y}}} \langle \boldsymbol{\alpha}, \boldsymbol{x}\rangle$ *is the support function of* $\mathcal{L}_0 I^*_{\boldsymbol{Y}}$.

Remark 5. Condition (H2) is an admissibility condition on sets A. Recall that $\mathcal{L}_0 I_{\boldsymbol{Y}}$ is the set of all the points $\boldsymbol{y}$ for which $I_{\boldsymbol{Y}}(\boldsymbol{y}) = 0$. Intuitively, these are the points of the typical behavior of $\boldsymbol{Y}$. Recall also that in the setting of Cramér's theorem, when $\boldsymbol{Y}_n = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_n$, for an i.i.d. sequence of random vectors $\{\boldsymbol{X}_n \colon n \in \mathbb{N}\}$ in $\mathbb{R}^d$, $\mathcal{L}_0 I_{\boldsymbol{Y}} = \{\mathbb{E}\boldsymbol{X}_1\}$. Therefore, the set $\mathcal{C}_\delta$ can be thought of as the δ-neighborhood of all such points of typical behavior of $\boldsymbol{Y}$, and, thus, A can be any general set that avoids the 'central tendency' $\text{cone}(\mathcal{C}_\delta) = \{\lambda\boldsymbol{x} \colon \lambda > 0,\ \boldsymbol{x} \in \mathcal{C}_\delta\}$ of $\boldsymbol{Y}$.

The straightforward major differences between the current approach and the one of Collamore (1996) are the following. Collamore (1996) considers the multidimensional case and sets satisfying (H2), but only allows linear scaling. The sequence $\boldsymbol{Y}$ has to satisfy (H1') and the sequence of measures corresponding to $\boldsymbol{Y}_n/n$, the LDP with *convex* rate function. In our setup we considered the multidimensional case and regularly varying scalings, general sets in the $d = 1$ case, but only quadrants $\{\boldsymbol{x} \in \mathbb{R}^d \colon \boldsymbol{x} > \boldsymbol{q}\}$, for any $\boldsymbol{q} > \boldsymbol{0}$, in the $d > 1$ case. Furthermore, we do not require the LDP to hold, but impose the restricted LDP hypothesis allowing for *continuous* rate functions. We have already explained that nonlinear scalings allow us to incorporate, for instance, long/short range dependent sequences stemming from, for example, fractional Brownian motion. Also, as explained in Theorem 1, it is possible that the restricted LDP holds when the LDP does not and vice versa.

Observe that, if $I_{\boldsymbol{Y}}$ is continuous where finite, then the restricted LDP holds with $J_{\boldsymbol{Y}}(\boldsymbol{q}) = \inf_{\boldsymbol{x}\geq\boldsymbol{q}} I_{\boldsymbol{Y}}(\boldsymbol{x})$. By Rockafellar (1970, Theorem 13.5), $\tilde{I}_{\boldsymbol{Y}}(\boldsymbol{x})$ is equal to the closure of $L(\boldsymbol{x}) = \inf_{\tau>0} \tau^{-1} I_{\boldsymbol{Y}}(\tau\boldsymbol{x})$, that is, the greatest lower semicontinuous function majorized by L. Furthermore, if $A = \{\boldsymbol{x} \in \mathbb{R}^d \colon \boldsymbol{x} > \boldsymbol{q}\}$, for some $\boldsymbol{q} > \boldsymbol{0}$, satisfies (H2) then the upper and the lower bound in Theorem 6 are equal to

$$\inf_{\boldsymbol{x}\geq\boldsymbol{q}} \tilde{I}_{\boldsymbol{Y}}(\boldsymbol{x}) = \inf_{\boldsymbol{x}\geq\boldsymbol{q}} \inf_{\tau>0} \tau^{-1} I_{\boldsymbol{Y}}(\tau\boldsymbol{x}) = \inf_{\tau>0} \tau^{-1} \inf_{\boldsymbol{x}\geq\boldsymbol{q}} I_{\boldsymbol{Y}}(\tau\boldsymbol{x}) = \inf_{\tau>0} \tau^{-1} J_{\boldsymbol{Y}}(\tau\boldsymbol{q}).$$

Hence, if, in addition, we are allowed to take $N_0 = 1$ then Theorem 6 coincides with Theorem 5 in the case of linear scaling and convex rate functions.

5. Examples

In this section we discuss some of the examples in which the theory of this paper can be applied. Firstly, we shall consider an example that does not fit in the framework of any of the previous literature. This is due to its multidimensional nature and the fact that the rate function appearing there is nonconvex. Secondly, we shall consider a multidimensional Gaussian example and we shall try to recover some previously known results.

Let us begin by discussing an extension of Theorem 5 from sequences $\{\boldsymbol{W}_n \colon n \in \mathbb{N}\}$ to stochastic processes $\{\boldsymbol{W}_t \colon t \in \mathbb{R}_+\}$. To this end, we formulate an additional necessary hypothesis.

Extension hypothesis. It holds that

$$\lim_{u\to\infty}\sup_{n\in\mathbb{N}}\frac{\log\mathbb{P}(\exists t\in(n,n+1]\colon \boldsymbol{W}_t>u\boldsymbol{q})}{h(u)}=\lim_{u\to\infty}\sup_{n\in\mathbb{N}}\frac{\log\mathbb{P}(\boldsymbol{W}_n>u\boldsymbol{q})}{h(u)}.$$

The above hypothesis was also introduced in Duffy *et al.* (2003) in the one-dimensional case. Therein, it is argued that under this hypothesis, for the $d=1$ case, Theorem 3 extends from sequences $\{W_n\colon n\in\mathbb{N}\}$ to processes $\{W_t\colon t\in\mathbb{R}_+\}$. It is straightforward to conclude from the proof of Theorem 5 that this is also the case if $d>1$.

5.1. Application to heavy-tailed processes

Let us consider processes of the type described in Duffy *et al.* (2003, Section 3.2). To this end, we first introduce the heavy-tailed distribution $\mathcal{H}$ by

$$\mathbb{P}(\mathcal{H}\geq x)=l(x)\mathrm{e}^{-v(x)},$$

where l is a slowly varying function and $v\in\mathcal{RV}(V)$ with $V\in(0,1)$.

Now consider a continuous time process $\{Y_t\colon t\in\mathbb{R}_+\}$ taking the values 0 and 1, with the times spent in the 0 and 1 states being a sequence of i.i.d. random variables with the same distribution as $\mathcal{W}$. For $c>0$, define a process $\{Z_t\colon t\in\mathbb{R}_+\}$ via

$$Z_t=\int_0^t (Y_s-c)\,\mathrm{d}s. \tag{13}$$

It was shown by Duffy and Sapozhnikov (2008) that the family of probability measures $\{\mu_t : t\in\mathbb{R}_+\}$, where μ_t is the law of Z_t/t, satisfies the LDP with speed function $v(t)=t^V$ and the (nonconvex!) rate function

$$I_c(x)=\begin{cases}(1-2(x+c))^V & \text{if } x\in[-c,\frac12-c],\\ (2(x+c)-1)^V & \text{if } x\in[\frac12-c,1-c],\\ \infty & \text{otherwise.}\end{cases} \tag{14}$$

Now consider two independent processes $\{Y_t^1 : t\in\mathbb{R}_+\}$ and $\{Y_t^2 : t\in\mathbb{R}_+\}$ defined as above and construct the corresponding processes $Z^1=\{Z_t^1\colon t\in\mathbb{R}_+\}$ and $Z^2=\{Z_t^2\colon t\in\mathbb{R}_+\}$ via (13) with constants $c_1>0$ and $c_2>0$, respectively. Now let $\boldsymbol{Z}=\{\boldsymbol{Z}_t\colon t\in\mathbb{R}_+\}$, where $\boldsymbol{Z}_t=(Z_t^1,Z_t^2)$. Note that the rate function $I_{\boldsymbol{Z}}$ corresponding to $\boldsymbol{Z}$ is given by $I_{\boldsymbol{Z}}(\boldsymbol{x})=I_{c_1}(x_1)+I_{c_2}(x_2)$, where I_{c_1} and I_{c_2} are given by (14). Finally, define a new process $\boldsymbol{W}=\{\boldsymbol{W}_t : t\in\mathbb{R}_+\}$ via $(W_t^1,W_t^2)=(Z_t^1,Z_t^1+Z_t^2)$. According to the contraction principle Dembo and Zeitouni (1998, Theorem 4.2.1), the family of probability measures $\{\mu_t^{\boldsymbol{W}} : t\in\mathbb{R}_+\}$ on $\mathbb{R}^2$, where $\mu_t^{\boldsymbol{W}}$ is the law of $\boldsymbol{W}_t/t$, satisfies the LDP with speed $v(t)=t^V$ and rate function $I_{\boldsymbol{W}}$ given by

$$I_{\boldsymbol{W}}(\boldsymbol{x})=\inf_{\boldsymbol{v}\in\mathbb{R}^2:(v_1,v_2)=(x_1,x_2-x_1)}(I_{c_1}(v_1)+I_{c_2}(v_2))=I_{c_1}(x_1)+I_{c_2}(x_2-x_1).$$

Hence, $\boldsymbol{W}$ satisfies the restricted LDP hypothesis with rate function $J_{\boldsymbol{W}}(\boldsymbol{q})=\inf_{\boldsymbol{x}\geq\boldsymbol{q}} I_{\boldsymbol{W}}(\boldsymbol{x})$, $\boldsymbol{q}\geq\boldsymbol{0}$.

Let $\hat{q}_1 = q_1 + c_1$ and $\hat{q}_2 = q_2 + c_1 + c_2$. Elementary calculus reveals that, with $\boldsymbol{q} \geq \boldsymbol{0}$,

$$J_{\boldsymbol{W}}(\boldsymbol{q}) = \begin{cases} J^1_{\boldsymbol{W}}(\boldsymbol{q}) \\ J^2_{\boldsymbol{W}}(\boldsymbol{q}) \\ J^3_{\boldsymbol{W}}(\boldsymbol{q}) \\ \infty \end{cases} = \begin{cases} I_{\boldsymbol{W}}(q_1, q_1 + \frac{1}{2} - c_2) \\ I_{\boldsymbol{W}}(q_2 + c_2 - \frac{1}{2}, q_2) \\ I_{\boldsymbol{W}}(1 - c_1, q_2) \\ \infty \end{cases}$$

$$= \begin{cases} I(\hat{q}_1) & \text{if } \hat{q}_1 \in (c_1, 1], \hat{q}_2 \in (c_1 + c_2, \frac{1}{2} + \hat{q}_1], \\ I(\hat{q}_2 - \frac{1}{2}) & \text{if } \hat{q}_1 \in (c_1, 1], \hat{q}_2 \in [\frac{1}{2} + \hat{q}_1, \frac{3}{2}], \\ 1 + I(\hat{q}_2 - 1) & \text{if } \hat{q}_1 \in (c_1, 1], \hat{q}_2 \in [\frac{3}{2}, 2], \\ \infty & \text{if } \hat{q}_1 > 1 \text{ or } \hat{q}_2 > 2, \end{cases}$$

where $I = I_0$. Recall that $c_1, c_1 \in (\frac{1}{2}, 1)$, thus, we do not define $J_{\boldsymbol{W}}(\boldsymbol{q})$ for $\hat{q}_1 \leq \frac{1}{2}, \hat{q}_2 \leq 1$. Note also that, if $c_1 + c_2 \in [\frac{3}{2}, 2)$, then $\hat{q}_2 \geq \frac{3}{2}$, so that the first two cases in the definition of $J_{\boldsymbol{W}}$ do not occur. Finally, we can easily check that $J_{\boldsymbol{W}}$ is continuous on the interior of the set where it is finite and that $J_{\boldsymbol{W}}(\boldsymbol{0}) > 0$. Hence, the stability and the continuity hypothesis holds. In Figure 1 we illustrate the definition of the $J_{\boldsymbol{W}}(\boldsymbol{q})$; $c_1 + c_2 < \frac{3}{2}$ case.

When the coordinates of $\boldsymbol{W}$ are bounded above (by $1 - c_1$ and $2 - c_1 - c_2$, respectively), the uniform decay rate hypothesis is satisfied. Note also that, for every n and $t \in (n, n+1]$, $\boldsymbol{W}_t = \boldsymbol{W}_n + \boldsymbol{R}_{n,t}$, where $\boldsymbol{R}_{n,t} = (R^1_{n,t}, R^1_{n,t})$ with both $R^1_{n,t} = \int_n^t (Y^1_s - c_1)\,\mathrm{d}s$ and $R^2_{n,t} = \int_n^t (Y^1_s + Y^2_s - c_1 - c_2)\,\mathrm{d}s$ bounded independently of t. This immediately implies the extension hypothesis too. Therefore, the extended version of Theorem 5 applies. It gives, for every $\boldsymbol{q} > \boldsymbol{0}$,

$$\lim_{u\to\infty} \frac{1}{u^V} \log \mathbb{P}\bigg(\exists t \in \mathbb{R}_+ : \int_0^t (Y^1_s - c_1)\,\mathrm{d}s > uq_1, \int_0^t (Y^1_s + Y^2_s - c_1 - c_2)\,\mathrm{d}s > uq_2\bigg) = -\inf_{t\geq 0} \frac{J_{\boldsymbol{W}}(t\boldsymbol{q})}{t^V}.$$

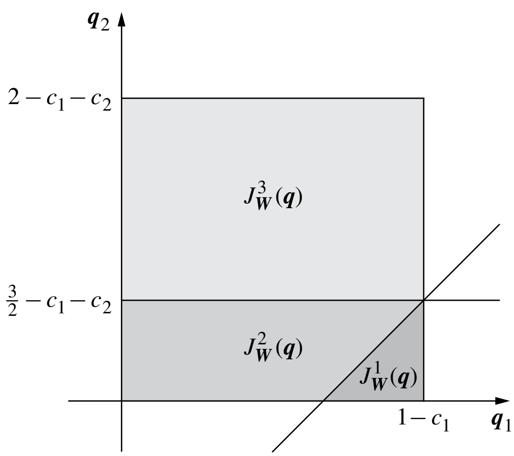

Figure 1: Definition of $J_{\boldsymbol{W}}(\boldsymbol{q})$; $c_1 + c_2 < \frac{3}{2}$ case. The rate function is ∞ for $q_1, q_2 > 0$ outside the shaded regions.

https://doi.org/10.1239/jap/1429282607 Published online by Cambridge University Press

Let $h(t) = J_W(t\boldsymbol{q})/t^V$, then with self-evident notation, if $c_1 + c_2 \in [\frac{3}{2}, 2)$,

$$\inf_{t\geq 0} h(t) = \begin{cases} h^3\left(\frac{1-c_1}{q_1}\right) \\ h^3\left(\frac{2-c_1-c_2}{q_2}\right) \end{cases} = \begin{cases} \left(\frac{q_1}{1-c_1}\right)^V \left(1 + \left[2\frac{q_2}{q_1}(1-c_1) + 2(c_1+c_2) - 3\right]^V\right), & q_1/q_2 \geq \frac{1-c_1}{2-c_1-c_2}, \\ 2\left(\frac{q_2}{2-c_1-c_2}\right)^V, & q_1/q_2 \leq \frac{1-c_1}{2-c_1-c_2}, \end{cases}$$

and if $c_1 + c_2 \in (1, \frac{3}{2})$,

$$\inf_{t\geq 0} h(t) = \begin{cases} h^1\left(\frac{1-c_1}{q_1}\right) \\ h^2\left(\frac{3/2-c_1-c_2}{q_2}\right) \end{cases} = \begin{cases} \left(\frac{q_1}{1-c_1}\right)^V, & q_1/q_2 \geq \frac{1-c_1}{3/2-c_1-c_2}, \\ \left(\frac{q_2}{3/2-c_1-c_2}\right)^V, & q_1/q_2 \leq \frac{1-c_1}{3/2-c_1-c_2}. \end{cases}$$

5.2. Application to Gaussian processes

In this subsection we consider an example from Dębicki *et al.* (2010, Section 3.2). That is, let $\boldsymbol{Y} = \{\boldsymbol{Y}(t) : t \in \mathbb{R}_+\}$ be a centered Gaussian process in $\mathbb{R}^d$ with stationary increments and covariance matrix $\Sigma_t = \operatorname{diag}(c_1\sigma^2(t), \dots, c_d\sigma^2(t))$, so that the coordinates of $\boldsymbol{Y}(t) = (Y_1(t), \dots, Y_d(t))$ are independent and, for each $i = 1, \dots, d$, $\mathbb{V}\mathrm{ar}(Y_i(t)) = c_i\sigma^2(t)$, for some $c_i > 0$, and $\sigma^2 \in \mathcal{RV}(\gamma)$, where $\gamma \in (0, 2)$; compare these assumptions with assumptions C1–C3 of Dębicki *et al.* (2010, Section 3.2)

For an invertible matrix S, define a new Gaussian process $\boldsymbol{W} = \{\boldsymbol{W}(t) : t \in \mathbb{R}_+\}$ in $\mathbb{R}^d$ via $\boldsymbol{W}(t) = S\boldsymbol{Y}(t) - \boldsymbol{i}(t)$, where $\boldsymbol{i} : \mathbb{R} \to \mathbb{R}^d$ is such that $\boldsymbol{i}(t) = (t, \dots, t)$. Set $a(t) = t$ and $v(t) = t^2\sigma^{-2}(t)$ and note that,

$$\Lambda_t(\boldsymbol{\alpha}) = \log \mathbb{E} \exp\left(\left\langle \boldsymbol{\alpha}, \frac{\boldsymbol{W}(t)}{a(t)} \right\rangle\right) = \frac{1}{2}\left\langle \boldsymbol{\alpha}, \frac{S\Sigma_t S^T}{t^2}\boldsymbol{\alpha} \right\rangle - \langle \boldsymbol{\alpha}, \boldsymbol{i}(1) \rangle,$$

so that the CGF variant of the uniform individual decay rate hypothesis holds with $F' = 2$. Furthermore,

$$\Lambda(\boldsymbol{\alpha}) = \lim_{t\to\infty} \frac{1}{v(t)} \Lambda_t(v(t)\boldsymbol{\alpha}) = \tfrac{1}{2}\langle \boldsymbol{\alpha}, SCS^T\boldsymbol{\alpha} \rangle - \langle \boldsymbol{\alpha}, \boldsymbol{i}(1) \rangle,$$

where $C = \operatorname{diag}(c_1, \dots, c_d)$, so that the CGF variant of the LDP hypothesis holds and the RLDP is satisfied with

$$J_W(\boldsymbol{q}) = \inf_{\boldsymbol{x}\geq \boldsymbol{q}} \Lambda^*(\boldsymbol{x}) = \tfrac{1}{2} \inf_{\boldsymbol{x}\geq \boldsymbol{q}} \langle S^{-1}(\boldsymbol{x} + \boldsymbol{i}(1)), C^{-1}S^{-1}(\boldsymbol{x} + \boldsymbol{i}(1)) \rangle,$$

where the form of Λ^* follows from Rockafellar (1970, Theorem 12.3). From this equation it follows that the stability and continuity hypothesis holds. Finally, by the stationarity of increments of $\boldsymbol{Y}$, for any $\varepsilon > 0$,

$$\mathbb{P}(\exists t \in (n, n+1] : \boldsymbol{W}(t) > u\boldsymbol{q}) \leq \mathbb{P}(\exists t \in [0, 1] : \boldsymbol{W}(t) > u\varepsilon\boldsymbol{q}) + \mathbb{P}(\boldsymbol{W}(n) > u(1-\varepsilon)\boldsymbol{q}). \quad (15)$$

Now, for any $\boldsymbol{x} > \boldsymbol{0}$, define a new Gaussian process via $Z(t) = \langle S\boldsymbol{Y}(t), \boldsymbol{x}\rangle / \langle \boldsymbol{x}, \varepsilon \boldsymbol{q}\rangle$. Using Borell's inequality (see, e.g. Adler (1990, Theorem 2.1)),

$$\begin{aligned}\log \mathbb{P}(\exists t \in [0,1]\colon \boldsymbol{W}(t) > u\varepsilon\boldsymbol{q}) &\le \log \mathbb{P}(\exists t \in [0,1]\colon Z(t) > u) \\ &\le -\frac{(u-\mu)^2}{2\sigma^2},\end{aligned} \tag{16}$$

where $\mu = \mathbb{E}\sup_{t\in[0,1]} Z(t)$ and $\sigma^2 = \sup_{t\in[0,1]} \mathbb{V}\mathrm{ar}(Z(t))$. Combining (15) and (16) we retrieve the extension hypothesis after proper optimization in $\varepsilon \to 0$. Hence, the extended version of Theorem 5 implies that,

$$\begin{aligned}&\lim_{u\to\infty} \frac{\sigma^2(u)}{u^2} \log \mathbb{P}(\exists t \in \mathbb{R}_+\colon \boldsymbol{W}(t) > u\boldsymbol{q}) \\ &\quad = -\frac{1}{2} \inf_{c\ge 0} \inf_{\boldsymbol{x}\ge c\boldsymbol{q}} \frac{\langle S^{-1}(\boldsymbol{x}+\boldsymbol{i}(1)), C^{-1}S^{-1}(\boldsymbol{x}+\boldsymbol{i}(1))\rangle}{c^{2-\gamma}}.\end{aligned} \tag{17}$$

From the change of variable $c \mapsto t^{-1}$, it follows that

$$\begin{aligned}&\inf_{c\ge 0} \inf_{\boldsymbol{x}\ge c\boldsymbol{q}} \frac{\langle S^{-1}(\boldsymbol{x}+\boldsymbol{i}(1)), C^{-1}S^{-1}(\boldsymbol{x}+\boldsymbol{i}(1))\rangle}{c^{2-\gamma}} \\ &\quad = \inf_{c\ge 0} \inf_{\boldsymbol{x}\ge \boldsymbol{q}} \frac{\langle S^{-1}(\boldsymbol{x}+\boldsymbol{i}(c^{-1})), C^{-1}S^{-1}(\boldsymbol{x}+\boldsymbol{i}(c^{-1}))\rangle}{c^{-\gamma}} \\ &\quad = \inf_{t\ge 0} \inf_{\boldsymbol{x}\ge \boldsymbol{q}} \frac{\langle S^{-1}(\boldsymbol{x}+\boldsymbol{i}(t)), C^{-1}S^{-1}(\boldsymbol{x}+\boldsymbol{i}(t))\rangle}{t^{\gamma}}.\end{aligned}$$

Therefore, (17) coincides with Dębicki *et al.* (2010, Proposition 2).

References

Adler, R. J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes* (Inst. Math. Statist. Lecture Notes Monogr. Ser. **12**). Institute of Mathematical Statistics, Hayward, CA.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press.

Collamore, J. F. (1996). Hitting probabilities and large deviations. *Ann. Prob.* **24,** 2065–2078.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.

Dębicki, K., Kosiński, K. M., Mandjes, M. and Rolski, T. (2010). Extremes of multidimensional Gaussian processes. *Stoch. Process. Appl.* **120,** 2289–2301.

Duffield, N. G. and O'Connell, N. (1995). Large deviations and overflow probabilities for the general single-server queue, with applications. *Math. Proc. Camb. Phil. Soc.* **118,** 363–374.

Duffy, K. R. and Sapozhnikov, A. (2008). The large deviation principle for the on–off Weibull sojourn process. *J. Appl. Prob.* **45,** 107–117.

Duffy, K., Lewis, J. T. and Sullivan, W. G. (2003). Logarithmic asymptotics for the supremum of a stochastic process. *Ann. Appl. Prob.* **13,** 430–445.

Glynn, P. W. and Whitt, W. (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. In *Studies in Applied Probability* (J. Appl. Prob. Spec. Vol. **31A**), Applied Probability Trust, Sheffield, pp. 131–156.

Kesidis, G., Walrand, J. and Chang, C.-S. (1993). Effective bandwidths for multiclass Markov fluids and other ATM sources. *IEEE/ACM Trans. Networking* **1,** 424–428.

Lelarge, M. (2008). Tail asymptotics for discrete event systems. *Discrete Event Dyn. Syst.* **18,** 563–584.

Ney, P. and Nummelin, E. (1987). Markov additive processes. I. Eigenvalue properties and limit theorems. *Ann. Prob.* **15,** 561–592.

Ney, P. E. and Robinson, S. M. (1995). Polyhedral approximation of convex sets with an application to large deviation probability theory. *J. Convex Anal.* **2,** 229–240.

Reich, E. (1958). On the integrodifferential equation of Takács. I. *Ann. Math. Statist.* **29,** 563–570.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.