ESAIM: PS 15 (2011) 83–109
DOI: 10.1051/ps/2009010

ESAIM: Probability and Statistics
www.esaim-ps.org

ON THE LARGE DEVIATIONS OF A CLASS OF MODULATED ADDITIVE PROCESSES

Ken R. Duffy[1], Claudio Macci[2] and Giovanni Luca Torrisi[3]

Abstract. We prove that the large deviation principle holds for a class of processes inspired by semi-Markov additive processes. For the processes we consider, the sojourn times in the phase process need not be independent and identically distributed. Moreover the state selection process need not be independent of the sojourn times. We assume that the phase process takes values in a finite set and that the order in which elements in the set, called states, are visited is selected stochastically. The sojourn times determine how long the phase process spends in a state once it has been selected. The main tool is a representation formula for the sample paths of the empirical laws of the phase process. Then, based on assumed joint large deviation behavior of the state selection and sojourn processes, we prove that the empirical laws of the phase process satisfy a sample path large deviation principle. From this large deviation principle, the large deviations behavior of a class of modulated additive processes is deduced. As an illustration of the utility of the general results, we provide an alternate proof of results for modulated Lévy processes. As a practical application of the results, we calculate the large deviation rate function for a processes that arises as the International Telecommunications Union's standardized stochastic model of two-way conversational speech.

Mathematics Subject Classification. 60F10, 60K15.

Received November 27, 2008. Revised May 8, 2009.

1. Introduction

Many stochastic models are most readily analyzed by expressing them as functions of well understood stochastic processes. Properties of the well understood processes can often be preserved by the representation. For example, the Large Deviation Principle (LDP) is preserved by quasi-continuous maps *via* the contraction principle [20] and weak convergence is preserved by continuous maps *via* the continuous mapping theorem [6].

In this article we illustrate how functional LDP techniques can be used to prove qualitative properties of a class of processes created through the apparently complex interactions of well behaved primitives. We demonstrate that the primary difficulty is in developing an appropriate sample path representation, whereupon the main results follow by the application of machinery that is, by now, standard. If additional independence assumptions

Keywords and phrases. Large deviations, modulated additive processes, speech models.

[1] Hamilton Institute, National University of Ireland, Maynooth, Ireland; ken.duffy@nuim.ie

[2] Dipartimento di Matematica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Rome, Italy; macci@mat.uniroma2.it

[3] Istituto per le Applicazioni del Calcolo "Mauro Picone", Consiglio Nazionale delle Ricerche (CNR), Via dei Taurini 19, 00185 Rome, Italy; torrisi@iac.rm.cnr.it

Article published by EDP Sciences

© EDP Sciences, SMAI 2011

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

hold, more quantitative conclusions can be drawn. We illustrate this by showing how results for modulated Lévy processes known by more traditional methods can be recovered by the functional LDP approach.

We illustrate the practical merits of the functional LDP approach by considering the application of our general results to a stochastic process defined in the ITU-T standard for modeling artificial conversation speech. These internationally standardized stochastic models of information sources are becoming increasingly common. For equipment to be standards-compliant, their performance is evaluated and bench-marked when processing work generated by these stochastic models. We calculate the large deviations rate function for ITU-T P.59, the stochastic model for a two way speech conversation. We use the rate function to deduce estimates on this processes queueing properties at a single server queue, which are used by teletraffic engineers as quality of service metrics.

2. Modulated processes

Modulated processes are constructed by concatenating together parts of paths of distinct processes to make a new process. Thus, at different times, the sample paths of the modulated process possess characteristics of the distinct underlying processes. The stochastic process that determines which underlying process is selected as a function of time is called the *phase process*.

The study of modulated processes dates back to at least the 1960s with Neveu's work on F-processes [30] and Çinlar's work on Markov Additive Processes (MAPs) [10] (which form a super-set of F-processes). Large deviation results for MAPs can be found in papers of Iscoe *et al.* [22], Ney and Nummelin [31,32] and Lehtonen and Nyrhinen [25]. Early applications of these processes used them successfully as models of packetized voice data, see, for example, Heffes and Lucantoni [21].

In MAPs the phase process is Markovian, so that underlying processes are followed for exponentially distributed periods of time, at the end of which the next process to be followed is chosen in a Markovian manner. The underlying processes are then required to have conditionally independent increments given the state of the phase process. Sometimes the underlying processes are required to be counting processes, as in recent work of Latouche *et al.* [23], but this is not always the case.

MAPs, and in particular their subclasses of batch Markov Arrival Processes and Markov Modulated Poisson Processes, have been widely adopted as tools in queueing theory as they enable the construction of traffic sources with bursty statistics. For examples of developments of results for these processes in queueing theory, see the papers by Pacheco and Prabhu [34] and Breuer [8].

Markov modulated processes are also used in finance and risk theory where the Markov state can be used to represent fundamental market conditions. For a recent example of work on Markov Modulated Lévy processes in risk theory see Asmussen and Pihlsgård [4], and for a recent example of work on finance with a Markov Modulated Drift process see Rieder and Bäuerle [37]. For an introduction to this area, see Asmussen [1,2].

Semi-Markov modulated (or semi-Markov additive) processes form a generalization to MAPs where the period of time that underlying stochastic processes are followed for are no longer exponential, but still form i.i.d. sequences; see *e.g.* Dshalalow and Russell [15], Özekici and Soyer [33] and Dshalalow [14]. The decision of which process to follow at the end of a period is again Markovian. As in the MAP setting, the underlying processes are assumed to be independent of the phase process.

In this paper we consider a natural collection of processes that are inspired by finite state semi-Markov additive processes. We relax the independence assumptions that are typically present within the construction of the phase process. This allows us to include in the treatment, for example, phase processes that visit states following a periodic structure or whose sojourn times are correlated.

We study the Large Deviation Principle (LDP) for a process that is modulated by a phase process that takes values in a finite state space $\{1, \ldots, K\}$. Our results give estimates on the likelihood that the sample paths of the empirical laws of the phase process behave atypically and estimates on the likelihood that the modulated additive process behaves atypically.

The phase process $\{Y_t, t \in \mathbb{R}_+\}$ is characterized by a state selection sequence $\{\sigma_n, n \in \mathbb{N}\}$, where $\sigma_n \in \{1, \ldots, K\}$ is the n^{th} state visited, and by sojourn times $\{\tau_n^k, n \in \mathbb{N}\}_{k=1}^K$, where τ_n^k is the time that the process

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

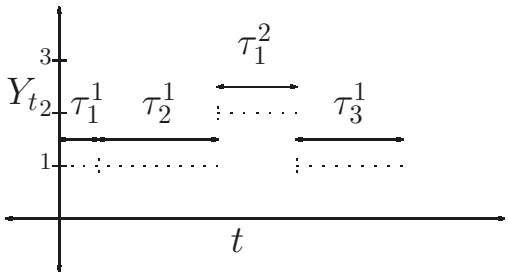


FIGURE 1. Example construction of the phase process $\{Y_t\}$.

spends on its n^{th} visit to state k. Throughout this paper we assume that all sojourn times τ_n^k are almost surely positive, $P(\tau_n^k > 0) = 1$. In general, we will not assume that the $\{\sigma_n\}$ and $\{\tau_n^k\}_{k=1}^K$ are independent of each other or that they are formed by i.i.d. random variables.

For a specific example of the construction of the phase process from $\{\sigma_n\}$ and $\{\tau_n^k\}_{k=1}^K$, consider the sequence $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 2, \sigma_4 = 1, \ldots$ The corresponding phase process would be

$$Y_t = \begin{cases} 1 & \text{if } 0 \le t < \tau_1^1 \\ 1 & \text{if } \tau_1^1 \le t < \tau_1^1 + \tau_2^1 \\ 2 & \text{if } \tau_1^1 + \tau_2^1 \le t < \tau_1^1 + \tau_2^1 + \tau_1^2 \\ 1 & \text{if } \tau_1^1 + \tau_2^1 + \tau_1^2 \le t < \tau_1^1 + \tau_2^1 + \tau_1^2 + \tau_3^1 \\ \vdots & \vdots \end{cases}$$

as illustrated in Figure 1.

We are interested in the large deviations of additive processes modulated by $\{Y_t\}$. That is, associated to each state $k \in \{1, \ldots, K\}$ there is an underlying stochastic process. At time t, the additive process consists of the sum over all K underlying processes, where for each state k, the process is evaluated at a time determined by how long Y_t has spent in state k by time t.

The rest of this paper is organized as follows. In Section 3 we construct a representation for the sample paths of the empirical laws of the phase process, $\{Y_t\}$, in terms of the sample paths of the empirical laws of the state selection process $\{\sigma_n\}$ and the sample paths of the partial sums of the sojourn times $\{\tau_n^k\}_{k=1}^K$ processes. This construction plays a crucial role in deriving the large deviation principles of this paper and is our primary contribution.

In Section 4 we recall the basic definitions of the theory of large deviations and, in Theorem 4.1, deduce a functional LDP for the empirical laws of the phase process $\{Y_t\}$ based on assumed sample path large deviation behavior for the empirical laws of the state selection process $\{\sigma_n\}$ and the partial sums of the sojourn times $\{\tau_n^k\}_{k=1}^K$.

We choose to work in the topology popularized by Ganesh *et al.* [17] on the space of continuous functions indexed by the positive real line. The topology is induced by a generalization of a supremum norm. By working in that space and its topology we are focusing on sojourn time processes that satisfy a uniform super-exponential tail condition, as for light tailed random variables it has been known since Mogulskii's [28] work that supremum-norm topologies are not appropriate for the sample path LDP. For light tailed sojourn times, one should use a generalization of the Skorohod topologies on the space of càdlàg functions indexed by non-compact sets, as in, for example, the work of Puhalskii [35] and Puhalskii and Whitt [36].

From the functional LDP for the empirical laws of the phase process, $\{Y_t\}$, we deduce the LDP for processes whose behavior is modulated by $\{Y_t\}$. In Section 5, Theorem 5.1 treats the large deviations of a process that accumulates a fixed value $v_k \in \mathbb{R}$ when $Y_t = k$. This process has applications in the modeling many of the standardized stochastic models of traffic processes in telecommunications. In particular, we give an application

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

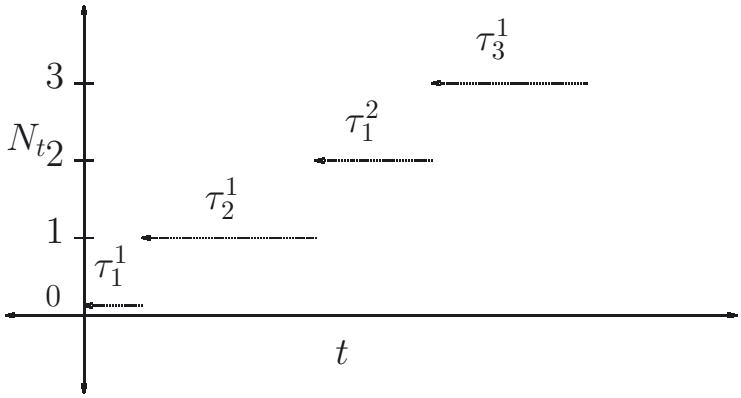


FIGURE 2. Example construction of the counting process $\{N_t\}$.

of the theory to the ITU-T standard for modeling artificial conversation speech. In practice, samples from this model are used to check that voice processing devices are standards compliant.

Results on overflow probabilities when traffic from modulated stochastic sources are fed into a single server queue are given in Section 6. Propositions 6.2 and 6.3 present new sufficient conditions under which these overflow probabilities decay asymptotically exponentially, demonstrating how this methodology can yield results of practical value. Finally, in Section 7 we consider modulated Lévy processes, which have applications in risk theory and finance, to illustrate how, under additional hypotheses, the general qualitative results can recover quantitative results. All proofs are deferred to Section 8.

3. A REPRESENTATION OF THE EMPIRICAL LAWS OF THE PHASE PROCESS

The main tool of this paper is the representation formula in Theorem 3.3. This formula represents the sample paths of the empirical laws of the phase process $\{Y_t\}$ in terms of the sample paths of the partial sums of the sojourn times $\{\tau_n^k\}$ and sample paths of the empirical laws of the state selection process $\{\sigma_n\}$.

To start with, we identify a tractable representation for Y_t in terms of the sequences $\{\sigma_n\}$ and $\{\tau_n^k\}_{k=1}^K$. We begin by introducing new random variables. With empty sums defined to be zero, for each $k \in \{1, \ldots, K\}$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$ define

$$L_n^k := \sum_{i=1}^{n} 1_{\{\sigma_i = k\}},\ S_n^k := \sum_{i=1}^{n} \tau_i^k,\ (T_0 := 0)\ T_n := \sum_{k=1}^{K} S_{L_n^k}^k,\ N_t := \sup\{n : T_n \le t\}. \tag{3.1}$$

The first object records the number of the first n states visited that are state k, so that $(L_n^1/n, \ldots, L_n^K/n)$ is the empirical law of $\{\sigma_1, \ldots, \sigma_n\}$. The second records the total time the phase process has spent in state k after it has been visited n times. The third records the total time that has passed when n states have been visited. The fourth records the number of state selections that have occurred before time t.

Corresponding to the example given in Section 2 ($\{Y_t\}$ in Fig. 1), Figure 2 plots N_t against t and Figure 3 plots the cumulative time spent in state 1 process, $S_{L_{N_t}^1}^1$, against t.

We can represent the phase process $\{Y_t\}$ at time t in terms of the random variables in equation (3.1):

$$Y_t = \sigma_{N_t+1}. \tag{3.2}$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

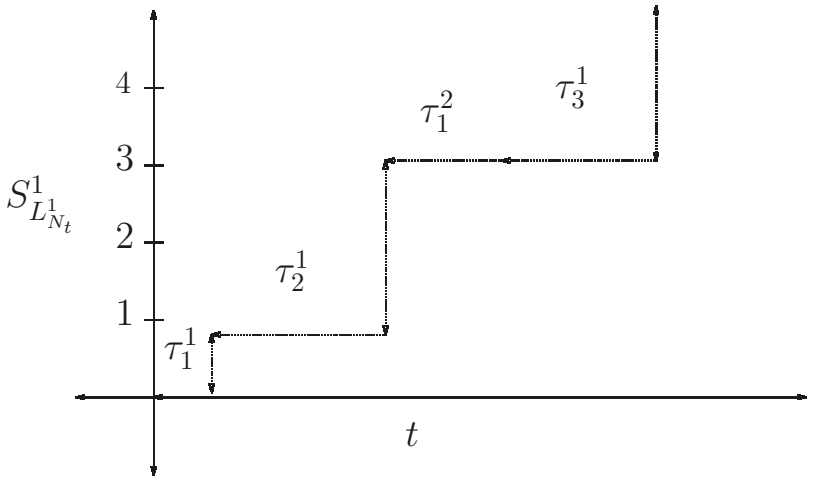


FIGURE 3. Example construction of the cumulative time spent in state 1 process $\{S^1_{L^1_{N_t}}\}$.

For each $k \in \{1, \ldots, K\}$, we define Z^k_t as the amount of time the process $\{Y_t\}$ has spent in state k by time t:

$$Z^k_t := S^k_{L^k_{N_t}} + (t - T_{N_t}) 1_{\{\sigma_{N_t+1}=k\}}. \tag{3.3}$$

Using equations (3.3) and (3.1) we have

$$\sum_{k=1}^{K} Z^k_t = \sum_{k=1}^{K} S^k_{L^k_{N_t}} + t - T_{N_t} = t.$$

The process $\{(Z^1_t/t, \ldots, Z^K_t/t)\}$ is the empirical measure process for $\{Y_t\}$. That is, Z^k_t/t records the fraction of time $\{Y_s, s \le t\}$ spent in the state k by time t.

We will construct a representation for the normalized sample paths for the empirical laws of $\{Y_t\}$ as a function of the normalized sample paths of the partial sums of $\{\sigma_n\}$ and $\{\tau^k_n\}^K_{k=1}$.

Remark. In most problems involving sample path representations it is natural to write the quantity of interest in terms of the càdlàg paths (paths which are right continuous with finite left limits), where the sample paths have discontinuities at the points where random variables are added. The use of continuous approximations to these paths, the polygonal sample paths, is typically a mathematical convenience and usually results in small errors that for large deviations must be managed through exponentially good approximations. However, here it will transpire that the continuous sample paths are the most natural building blocks and no exponential approximations will be necessary.

Throughout this paper we denote the integer part of $x \in \mathbb{R}$ by $[x]$. For a stochastic process indexed by the integers, $\{W_n, n \in \mathbb{N}\}$, we define its polygonal sample path to be

$$W_1(t) = W_{[t]} + (t - [t]) \left(W_{[t]+1} - W_{[t]}\right) \qquad \text{for all } t \in \mathbb{R}_+.$$

Its normalized polygonal sample paths are then defined for each $n \in \mathbb{N}$ to be

$$W_n(t) := \frac{1}{n} W_1(nt) = \frac{1}{n} W_{[nt]} + \left(t - \frac{[nt]}{n}\right) \left(W_{[nt]+1} - W_{[nt]}\right) \qquad t \in \mathbb{R}_+.$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

In what follows we consider the normalized polygonal sample paths for the sequences $\{L_n^k\}_{k=1}^K$, $\{S_n^k\}_{k=1}^K$ and $\{T_n\}$. Define

$$k^* = k^*(n,t) := \sigma_{[nt]+1}, \qquad (n \in \mathbb{N}, t \in \mathbb{R}_+) \tag{3.4}$$

which is the $([nt]+1)^{\text{th}}$ state visited by the process. We then have for all $n \in \mathbb{N}$, $t \in \mathbb{R}_+$,

$$L_n^k(t) := \frac{1}{n} L_{[nt]}^k + \left(t - \frac{[nt]}{n}\right) 1_{\{k=k^*\}}, \tag{3.5}$$

$$S_n^k(t) := \frac{1}{n} S_{[nt]}^k + \left(t - \frac{[nt]}{n}\right) \tau_{[nt]+1}^k \tag{3.6}$$

and

$$T_n(t) := \frac{1}{n} T_{[nt]} + \left(t - \frac{[nt]}{n}\right) \tau_{L_{[nt]}^{k^*}+1}^{k^*}. \tag{3.7}$$

The sample path $L_n^k(\cdot)$ is for the empirical law for the number of visits to state k, while $S_n^k(\cdot)$ is the sample path for the total time spent in state k. These are the basic paths from which we will deduce all the others. The sample path $T_n(\cdot)$ is for the total time that has passed after a given number of states have been visited.

The following lemma gives a representation of $T_n(\cdot)$ in terms of the sample paths $S_n^1(\cdot), \ldots, S_n^K(\cdot), L_n^1(\cdot), \ldots, L_n^K(\cdot)$.

Lemma 3.1 ($S \circ L$ representation)**.** *The following equation holds*

$$S_n^k(L_n^k(t)) = \frac{1}{n} S_{L_{[nt]}^k}^k + \left(t - \frac{[nt]}{n}\right) 1_{\{k=k^*\}} \tau_{L_{[nt]}^k+1}^k \tag{3.8}$$

and, as a consequence, we have the following representation for $T_n(\cdot)$ in terms of the sample paths $S_n^k(\cdot)$ and $L_n^k(\cdot)$:

$$T_n(\cdot) = \sum_{k=1}^K S_n^k \circ L_n^k(\cdot).$$

Next we need to introduce the sample paths associated with the counting process $\{N_t\}$. Define

$$k^{**} = k^{**}(n,t) := \sigma_{N_{nt}+1}, \tag{3.9}$$

which, by equation (3.2), is the phase process Y_{nt} at time nt. Then we define the normalized sample paths

$$N_n(t) := \frac{1}{n} N_{nt} + \left(t - \frac{T_{N_{nt}}}{n}\right) \left(\tau_{L_{N_{nt}}^{k^{**}}+1}^{k^{**}}\right)^{-1}, \qquad \text{for } n \in \mathbb{N}, t \in \mathbb{R}_+. \tag{3.10}$$

Considering $N_1(\cdot)$, it is clear that it is the natural polygonal approximation to the sample path of $\{N_t\}$, whose càdlàg paths would have discontinuities at not-necessarily integer times. It records the number of states that have been visited by time t, plus a linear proportion of the time that has passed prior to the following state-change. It therefore takes values in $\mathbb{R}_+$ not in $\mathbb{N}$.

Given two nonnegative stochastic processes with strictly increasing trajectories $\{W_1(t),\, t \in \mathbb{R}_+\}$ and $\{W_2(t),\, t \in \mathbb{R}_+\}$, we say that they are *inverse* to each other if $W_1(W_2(t)) = W_2(W_1(t)) = t$ for all $t \in \mathbb{R}_+$, and we write $W_1^{-1}(\cdot) = W_2(\cdot)$ and $W_2^{-1}(\cdot) = W_1(\cdot)$.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Lemma 3.2 ($N_n = T_n^{-1}$)**.** *For all $n \in \mathbb{N}$, $T_n(\cdot)$ and $N_n(\cdot)$ are inverse to each other.*

We define the normalized polygonal sample path for the empirical laws for the time spent in state k, $\{Z_t^k\}$ defined in equation (3.3), to be

$$Z_n^k(t) := \frac{1}{n} S^k_{L^k_{N_{nt}}} + \left(t - \frac{T_{N_{nt}}}{n}\right) 1_{\{k=k^{**}\}}. \tag{3.11}$$

The following theorem gives a representation for $Z_n^k(\cdot)$ in terms of $S_n^k(\cdot)$ and $L_n^k(\cdot)$, $k \in \{1, \ldots, K\}$.

Theorem 3.3 ($Z_n^k = S_n^k \circ L_n^k \circ N_n$)**.** *We have the following representation of the normalized sample paths for the empirical laws $\{Z_n^k(\cdot)\}$ of the phase process $\{Y_t\}$ in terms of the normalized sample paths for the empirical laws $\{L_n^k(\cdot)\}$ of the state selection process $\{\sigma_n\}$ and the normalized sample paths $\{S_n^k(\cdot)\}$ for the partial sums of the sojourn times $\{\tau_n^k\}_{k=1}^K$. For all $n \geq 1$,*

$$Z_n^k(\cdot) = S_n^k \circ L_n^k \circ N_n(\cdot) = S_n^k \circ L_n^k \circ \left(\sum_{k=1}^K S_n^k \circ L_n^k\right)^{-1}(\cdot). \tag{3.12}$$

In the next section we will assume the LDP for $\{\left(S_n^1(\cdot), \ldots, S_n^K(\cdot), L_n^1(\cdot), \ldots, L_n^K(\cdot)\right), n \in \mathbb{N}\}$. This is a natural assumption as we can appeal to a general theorem that asserts this is the case for a large collection of processes. We will then deduce a functional LDP for $\{\left(Z_n^1(\cdot), \ldots, Z_n^K(\cdot)\right), n \in \mathbb{N}\}$ and an LDP for $\{\left(Z_n^1/n, \ldots, Z_n^K/n\right), n \in \mathbb{N}\}$ from the representation in equation (3.12).

4. Large deviations and the empirical laws of the phase process

For convenience we recall the definition of the Large Deviation Principle (LDP), which can be found in a standard text such as Dembo and Zeitouni [12], and introduce the function spaces we will use. Let $\mathcal{X}$ be a Hausdorff space with Borel σ-algebra $\mathcal{B}$ and let $\{\mu_n, n \in \mathbb{N}\}$ be a sequence of probability measures on $(\mathcal{X}, \mathcal{B})$. We say that $\{\mu_n, n \in \mathbb{N}\}$ satisfies the LDP in $\mathcal{X}$ with rate function $I : \mathcal{X} \to [0, +\infty]$ if I is lower semi-continuous,

$$-\inf_{x \in G} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mu_n[G] \qquad \text{and} \qquad \limsup_{n\to\infty} \frac{1}{n} \log \mu_n[F] \leq -\inf_{x \in F} I(x)$$

for all open G and all closed F. Rate functions whose level-sets $\{x : I(x) \leq \alpha\}$ are compact for all $\alpha \geq 0$ are said to be *good*. We say that a process $\{W_n, n \in \mathbb{N}\}$ satisfies the LDP if W_n is a realization of μ_n for each n. The main large deviation tool used throughout this paper is the Contraction Principle (*e.g.* Thm. 4.2.1 of [12]).

Let $\mathcal{C}[0, \infty)$ denote the collection of $\mathbb{R}$-valued continuous functions ϕ on $[0, \infty)$ such that $\phi(0) = 0$. Let $\mathcal{A}[0, \infty)$ denote the subset of $\mathcal{C}[0, \infty)$ whose elements are the integrals of Lebesgue integrable functions on $[0, x)$ for all $x > 0$. Define the space

$$\mathcal{Y} := \left\{\phi \in \mathcal{C}[0, \infty) : \lim_{t\to\infty} \frac{\phi(t)}{1+t} \text{ exists in } \mathbb{R}\right\}$$

and equip it with the topology induced by the norm

$$||\phi|| = \sup_{t \geq 0} \left|\frac{\phi(t)}{1+t}\right|. \tag{4.1}$$

Define

$$\mathcal{Y}_\uparrow := \{\phi \in \mathcal{Y} : \phi \text{ is strictly increasing and } \lim \phi(t)/(1+t) > 0\}$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

and

$$\mathcal{Y}_{\text{emp.meas.}} = \left\{ (\psi^1, \ldots, \psi^K) \in \mathcal{Y}^K : \text{each } \psi^k \text{ is non-decreasing and } \sum_{k=1}^{K} \psi^k(t) = t \right\},$$

Here $\mathcal{Y}_\uparrow$ is a metric subspace of $\mathcal{Y}$, and $\mathcal{Y}_{\text{emp.meas.}}$ is a metric subspace of $\mathcal{Y}^K$ equipped with the product topology. Moreover, $\mathcal{Y}_{\text{emp.meas.}}$ is closed.

Our main assumption is the following:

Assumption 4.1. *The sequence* $\{\left(S_n^1(\cdot), \ldots, S_n^K(\cdot), L_n^1(\cdot), \ldots, L_n^K(\cdot)\right), n \in \mathbb{N}\}$ *satisfies the LDP in* $(\mathcal{Y}_\uparrow)^K \times \mathcal{Y}_{\text{emp.meas.}}$ *with good rate function* I.

Assumption 4.1 holds for a large class of processes. Generalizing Mogulskii's theorem [28] for i.i.d. processes, Theorem 2 of Dembo and Zajic [11] establishes that, for a class of processes which satisfy a mixing condition and a uniform super-exponential tail condition, the polygonal sample paths of their partial sums satisfy the LDP in the space of continuous functions on $[0, 1]$ equipped with the supremum norm. Ganesh and O'Connell [18]* prove that these LDPs can be strengthened to hold in $\mathcal{Y}$ with the topology induced by the norm defined in equation (4.1). This is why it is reasonable to start with assumed large deviation properties for the sample paths of partial sums of $\{\sigma_n\}$ and $\{\tau_n^k\}_{k=1}^K$.

Theorem 4.1 (LDP for the empirical laws of the phase process)**.** *Under Assumption 4.1 we have:*

(1) *the normalized sample paths of the empirical laws of the phase process,* $\{\left(Z_n^1(\cdot), \ldots, Z_n^K(\cdot)\right), n \in \mathbb{N}\}$, *satisfy the LDP in* $\mathcal{Y}_{\text{emp.meas.}}$ *with good rate function*

$$J_{Z(\bullet)}(\eta^1, \ldots, \eta^K) = \inf \left\{ I(\phi^1, \ldots, \psi^K) : \phi^k \circ \psi^k \circ \left(\sum_{j=1}^{K} \phi^j \circ \psi^j \right)^{-1} = \eta^k, \ k \in \{1, \ldots, K\} \right\};$$

(2) *the empirical laws of the phase process,* $\{\left(Z_n^1/n, \ldots, Z_n^K/n\right), n \in \mathbb{N}\}$, *satisfy the LDP in* $\mathbb{R}^K$ *with good rate function*

$$J_{\mathbf{Z}}(z^1, \ldots, z^K) = \inf \left\{ I(\phi^1, \ldots, \psi^K) : \phi^k \circ \psi^k \circ \left(\sum_{j=1}^{K} \phi^j \circ \psi^j \right)^{-1} (1) = z^k, \ k \in \{1, \ldots, K\} \right\}. \tag{4.2}$$

Theorem 4.1 establishes that the LDP holds for the sample paths of the empirical laws of a large class of finite state phase processes. From this LDP, we can deduce LDPs for associated modulated additive processes, as we illustrate in the following sections.

5. Fluid processes: large deviations

Consider the setting where there is a fixed value $v_k \in \mathbb{R}$ associated with each state $k \in \{1, \ldots, K\}$. Define the modulated additive process by

$$X_t := \int_0^t v_{Y_s} \, \mathrm{d}s \qquad \text{for all } t \in \mathbb{R}_+. \tag{5.1}$$

*These results can also be found in Section 6 of [17].

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Processes such as these arise in models of traffic-sources in telecommunication networks where X_t represents the total amount of traffic generated by a source by time t. The transmission rates, $\{v_k\}_{k=1}^K$, depend upon the underlying application and change dynamically (see *e.g.* Markopoulou *et al.* [27]).

It will be helpful to use the following alternative representation to equation (5.1) for X_t:

$$X_t = \sum_{k=1}^{K} v_k Z_t^k = \sum_{k=1}^{K} v_k S^k_{L^k_{N_t}} + v_{\sigma_{N_t+1}}(t - T_{N_t}). \tag{5.2}$$

We define the normalized polygonal sample paths for $\{X_t\}$ as being scaled down by a factor n and speeded up by factor n,

$$X_n(t) := \frac{1}{n}\sum_{k=1}^{K} v_k S^k_{L^k_{N_{nt}}} + v_{\sigma_{N_{nt}+1}}\left(t - \frac{1}{n}T_{N_{nt}}\right) = \sum_{k=1}^{K} v_k Z_n^k(t). \tag{5.3}$$

Note that, for example, evaluating $X_n(1)$ results in the same expression as using equation (5.2) to determine X_n/n.

Theorem 5.1 (LDP for fluid processes)**.** *Under Assumption 4.1,* $\{X_n/n, n \in \mathbb{N}\}$ *satisfies the LDP in* $\mathbb{R}$ *with good rate function*

$$J_X(x) = \inf\left\{ I(\phi^1, \ldots, \psi^K) : \sum_{k=1}^{K} v_k\, \phi^k \circ \psi^k \circ \left(\sum_{j=1}^{K} \phi^j \circ \psi^j\right)^{-1}(1) = x \right\}. \tag{5.4}$$

Before identifying sufficient conditions under which equation (5.4) reduces to finite dimensional optimization, we make two elementary comments.

Firstly, define the set

$$\Delta_K := \left\{ (y_1, \ldots, y_K) \in [0,1]^K : \sum_{k=1}^{K} y_k = 1 \right\}. \tag{5.5}$$

By Lemma 4.1.5 [12] easily follows that $J_X(x) = \infty$ for $x \notin [\min_k v_k, \max_k v_k]$ and $J_{\mathbf{Z}}(z) = \infty$ for $z \notin \Delta_K$. Secondly, note that by Theorem 5.1, $\{X_t/t, t \in \mathbb{R}_+\}$ satisfies the LDP with good rate function J_X, indeed a straightforward computation shows that the processes $\{X_t/t, t \in \mathbb{R}_+\}$ and $\{X_{[t]}/[t], t \in \mathbb{R}_+\}$ are exponentially equivalent (see *e.g.* [12], Def. 4.2.10 and Thm. 4.2.13).

Without additional assumptions, the variational problem in equation (5.4) does not simplify further. This is typically the case in functional large deviations if the representation formula is involved. We present two sets of additional independence conditions in which it simplifies to a finite dimensional convex optimization problem.

5.1. Independent state selection and sojourn time processes

Firstly we consider the case where the state selection sequence $\{\sigma_n\}$ is independent of the sojourn times $\{\tau_n^k\}_{k=1}^K$, and $\{\tau_n^i\}$ is independent of $\{\tau_n^j\}$ for all $i \neq j$. However, the state selection sequence $\{\sigma_n\}$ need not consist of independent random variables, nor for each k need $\{\tau_n^k\}$ consist of independent random variables.

In the telecommunications application, this corresponds to the transmission rate being chosen independently of how long the transmission will take. However there can be correlations in how long a transmission rate is adopted once it has been selected. There may also be correlations within the state selection process.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Define the set

$$\mathcal{Y}^{\#} := \{(\phi^1, \dots, \psi^K) \in (\mathcal{Y}_{\uparrow})^K \times \mathcal{Y}_{\text{emp.meas.}} : \phi^1, \dots, \psi^K \in \mathcal{A}[0,\infty)\},$$

and consider the following assumption:

Assumption 5.1. *Assumption 4.1 holds and, in addition,*

$$I(\phi^1, \dots, \psi^K) = \begin{cases} \int_0^\infty \left(\sum_{k=1}^K I^k(\dot{\phi}^k(s)) + H(\dot{\psi}^1(s), \dots, \dot{\psi}^K(s)) \right) \mathrm{d}s, & \text{if } (\phi^1, \dots, \psi^K) \in \mathcal{Y}^{\#} \\ \infty & \text{otherwise,} \end{cases} \tag{5.6}$$

where $I^k : \mathbb{R}_+ \to [0,\infty]$ and $H : \Delta_K \to [0,\infty]$, with Δ_K defined in equation (5.5), are convex lower semi-continuous functions.

With this assumption in force, the functional infimization in equation (5.4) simplifies into a convex programme that can be readily solved numerically.

Corollary 5.2 (Independent state selection and sojourn times)**.** *Under Assumption 5.1, J_X is convex and given by*

$$J_X(x) = \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf_{x_1,\dots,x_K>0} \left\{ \sum_{k=1}^K y_k I^k\left(\frac{x_k}{y_k}\right) + yH\left(\frac{y_1}{y}, \dots, \frac{y_K}{y}\right) : \sum_{k=1}^K y_k = y, \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\}. \tag{5.7}$$

In the proof of Corollary 5.2 we point out that there exist $m_1, \dots, m_K \in (0,\infty)$ and $(l_1, \dots, l_K) \in \Delta_K$ such that $I^k(m_k) = 0$ $(k \in \{1, \dots, K\})$ and $H(l_1, \dots, l_K) = 0$. Now consider the following choice of the variables $y, y_1, \dots, y_K, x_1, \dots, x_K$ in (5.7): for $k \in \{1, \dots, K\}$ we set $y_k/y = l_k$ and $x_k/y_k = m_k$, whence we obtain $x_k = m_k l_k y$. Therefore we have $y = (\sum_k m_k l_k)^{-1}$ since $\sum_k x_k = 1$, and finally the constraint $\sum_k v_k x_k = x$ gives $x = \sum_k v_k m_k l_k / \sum_k m_k l_k$. Thus $J_X(x) = 0$ for $x = \sum_k v_k m_k l_k / \sum_k m_k l_k$.

Example. We illustrate the result in Corollary 5.2 by treating a process which arises in a telecommunications standard. We consider the large deviations of the total volume of traffic generated by a two way conversation.

The International Telecommunication Union (ITU), which has been a specialized agency of the United Nations since 1947, has a permanent organ called the ITU Standardization Sector (ITU-T). It is responsible for issuing "Recommendations" with a view to standardizing telecommunications on a worldwide basis. The ITU-T Recommendations are, effectively, worldwide standards.

In its Recommendation P.59 [39], which was approved in March 1993, the ITU-T gives a model for two way voice traffic. The Recommendation is based on papers such as those by Brady [7] and by Lee and Un [24]. The P.59 model for generating artificial conversational speech has three states: state 1 corresponds to mutual silence; state 2 corresponds to single talk; and state 3 corresponds to double talk. Conditioned on the state, the model's sojourn times are i.i.d. random variables with given means. We model the sojourn times as a deterministic pause plus a truncated exponential distribution. The state selection process, $\{\sigma_n\}$, forms a Markov chain with transition matrix

$$\pi = \begin{pmatrix} 0 & 1 & 0 \\ \alpha & 0 & 1-\alpha \\ 0 & 1 & 0 \end{pmatrix}, \tag{5.8}$$

where, in the standard, $\alpha = 0.4$. That is, both mutual silence and double talk are always followed by single talk. Single talk is followed by either mutual silence, with probability 0.4, or double talk, with probability 0.6.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

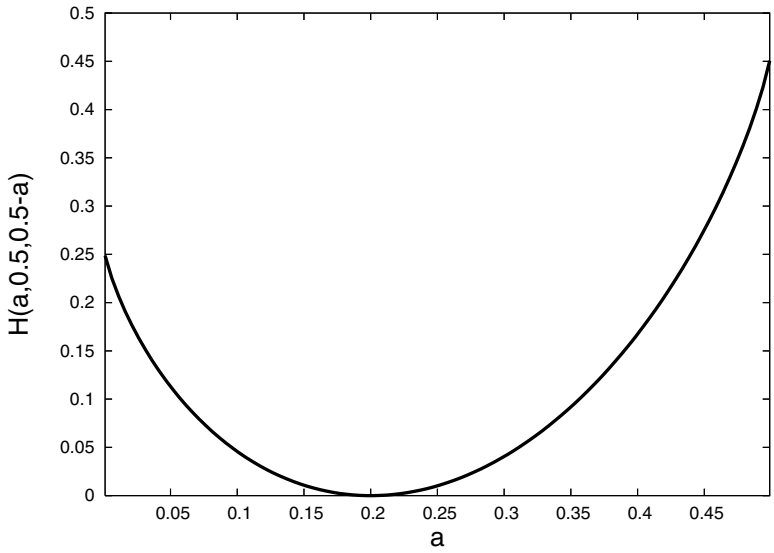


FIGURE 4. Empirical law rate function $H(a, 0.5, 0.5-a)$ for state selection in ITU-T P.59

Using equation (4.1.37) in Deuschel and Stroock [13], it is possible to calculate $H(\cdot,\cdot,\cdot)$ in equation (5.7) for this Markov chain. The rate function $H(x,y,z)$ is only finite if $(x,y,z) = (a, 1/2, 1/2-a)$, where $a \in [0, 1/2]$, and then

$$H\left(a, \frac{1}{2}, \frac{1}{2}-a\right) = \begin{cases} -\frac{1}{2}\log(1-\alpha) & \text{if } a = 0, \\ -(\frac{1}{2}-a)\log(1-\alpha) + (\frac{1}{2}-a)\log(1-2a) + a\log\left(\frac{2a}{\alpha}\right) & \text{if } a \in (0, 1/2), \\ -\frac{1}{2}\log(\alpha) & \text{if } a = 1/2. \end{cases}$$

With $\alpha = 0.4$, as in the standard, the rate function $H(a, 0.5, 0.5-a)$ is plotted in Figure 4.

Returning to the sojourn times, each τ_1^i has distribution

$$P(\tau > t) = \begin{cases} 1 & \text{if } t \le T_l \\ \exp(-\lambda_i t) & \text{if } t \in [T_l, T_u) \\ 0 & \text{if } t \ge T_u \end{cases}$$

where $0 < T_l < T_u$. In the standard $T_l = 200$ ms. We choose large T_u and then the $\{\lambda_i\}$ are selected so that the mean sojourn times are as in the Recommendation. This gives $\lambda_1 = 0.8555$, $\lambda_2 = 0.2261$ and $\lambda_3 = 0.4564$. Then $I^i(x) = \sup_{\theta\in\mathbb{R}}\left\{x\theta - \log \mathbb{E}[\exp(\theta\tau_1^i)]\right\}$. For the truncated exponential, this is a transcendental equation for $I^i(x)$ that can be readily solved numerically.

As mutual silence generates no traffic, we set $v_1 = 0$. With $v_2 = 1$, we let $v_3 = 2$ as double talk generates twice as much traffic as single talk. As H is infinite for a large range of values, J_X defined in equation (5.7) reduces to

$$J_X(x) = \inf_{y>0}\inf_{a\in[0,1/2]}\inf_{z\in(0,x)}\left\{ayI^1\left(\frac{2-x-z}{2ay}\right) + \frac{y}{2}I^2\left(\frac{2z}{y}\right) + \left(\frac{1}{2}-a\right)yI^3\left(\frac{x-z}{(1-2a)y}\right) + yH\left(a, \frac{1}{2}, \frac{1}{2}-a\right)\right\}. \tag{5.9}$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

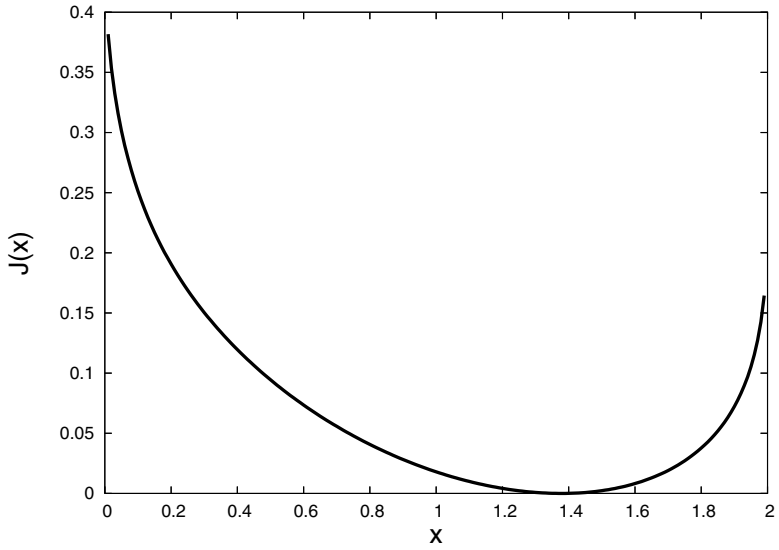


FIGURE 5. $J_X(x)$ vs. x for a two way voice conversation based on ITU-T P.59

This is a finite dimensional convex programme that is readily solved numerically. For ITU-T P.59, the rate function $J_X(x)$ is plotted in Figure 5. As $\alpha = 0.4$, the average volume of traffic produced is greater than 1, the amount caused by single talk. However, the rate function is skewed as when mutual silence periods occur they are likely to last for longer than double talk periods. Thus, using $J_X(\cdot)$ we can estimate the likelihood that a long lived conversation will generate more (or less) traffic than is expected.

Before concluding this subsection, we make one last remark that will prove useful in the consideration of the example in Section 7.

Remark. Under Assumption 5.1 the functional infimization in equation (4.2) for $J_{\mathbf{Z}}$ also simplifies into a convex programme that can be readily solved numerically. More precisely, arguing as in the proof of Corollary 5.2, for all $(z_1, \ldots, z_K) \in \Delta_K$,

$$J_{\mathbf{Z}}(z_1, \ldots, z_K) = \inf_{y_1, \ldots, y_K > 0} \left\{ \sum_{k=1}^{K} y_k I^k \left(\frac{z_k}{y_k} \right) + \left(\sum_{k=1}^{K} y_k \right) H \left(\frac{y_1}{\sum_k y_k}, \ldots, \frac{y_K}{\sum_k y_k} \right) \right\},$$

and $J_{\mathbf{Z}}(z_1, \ldots, z_K) = +\infty$ if $(z_1, \ldots, z_K) \notin \Delta_K$.

In anticipation of the example in Section 7 we remark that, if we assume that there exists $m > 0$ such that

$$I^k(x) = \begin{cases} 0 & \text{if } x = m \\ \infty & \text{if } x \neq m \end{cases} \quad \text{for all } k \in \{1, \ldots, K\},$$

then

$$J_{\mathbf{Z}}(z^1, \ldots, z^K) = \begin{cases} m^{-1} H(z^1, \ldots, z^K) & \text{if } (z^1, \ldots, z^K) \in \Delta_K \\ \infty & \text{otherwise.} \end{cases}$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

5.2. **Equally visited states**

A simpler setting is where all states are visited equally frequently. Define the set

$$\mathcal{Y}^* := \{(\phi^1, \dots, \psi^K) \in (\mathcal{Y}_\uparrow)^K \times \mathcal{Y}_{\text{emp.meas.}} : \phi^1, \dots, \phi^K \in \mathcal{A}[0,\infty), \psi^1(t) = \cdots = \psi^K(t) = t/K\},$$

and consider the following assumption:

Assumption 5.2. *Assumption 4.1 holds with*

$$I(\phi^1, \dots, \psi^K) = \begin{cases} \int_0^\infty I'(\dot\phi^1(s), \dots, \dot\phi^K(s))\mathrm{d}s & \text{if } (\phi^1, \dots, \psi^K) \in \mathcal{Y}^* \\ \infty & \text{otherwise,} \end{cases}$$

where $I' : \mathbb{R}_+^K \to [0,\infty]$ is a convex lower semi-continuous function.

Corollary 5.3 (equally visited states)**.** *Under Assumption 5.2, J_X is convex and given by*

$$J_X(x) = \inf_{y>0} \inf_{x_1,\dots,x_K>0} \left\{ y I'\left(\frac{x_1}{y}, \dots, \frac{x_K}{y}\right) : \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\}.$$

Example. In the case of $K = 2$ with $v_1 = 0$ and $v_2 = 1$, this gives the rate function for an alternating two-state on–off process in terms of the joint rate function for its sojourn times:

$$J_X(x) = \inf_{y>0} y\, I'\left(\frac{1-x}{y}, \frac{x}{y}\right). \tag{5.10}$$

A result related to that in equation (5.10), but for sub-exponential sojourn times, appears in Duffy and Sapozhnikov [16]. There the sojourn times are i.i.d. with Weibull distribution, $P(\tau > t) = \exp(-t^\alpha)$ with $\alpha \in (0,1)$, and the result is proved by non-functional techniques. However, it is interesting to note that the resulting rate function has the form $J_X(x) = \inf_y y^\alpha (I_W((1-x)/y) + I_W(x/y))$, where $I_W(\cdot)$ is the rate function for partial sums of i.i.d. Weibull distributed random variables which can be found in, *e.g.*, Nagaev [29] or Gantert [19].

6. Fluid processes: overflow probabilities

Define the random variable

$$X_{\sup} := \sup_{t>0}\{X_t - t\}.$$

As $X_{\sup}/n = \sup_{t>0}\{X_n(t) - t\}$, by proving the LDP for $\{\sup_{t>0}\{X_n(t) - t\}, n \in \mathbb{N}\}$ we are proving that $\{X_{\sup}/n, n \in \mathbb{N}\}$ satisfies the LDP. From this we can deduce that the tail of the distribution of $X_{\sup}$ is approximately exponential (see *e.g.* [17]). If we consider $\{X_t\}$ as representing a source of traffic that sends data at constant rate v_k when $Y_t = k$, then $X_{\sup}$ corresponds to the waiting time at a single server queue served at constant rate 1 (see *e.g.* Asmussen [3]).

We start with the following assumption:

Assumption 6.1. *Assumption 4.1 holds, and $\mathcal{Z}_{J_X} \subset (-\infty, 1)$ where $\mathcal{Z}_{J_X} := \{x \in \mathbb{R} : J_X(x) = 0\}$ and J_X is the rate function in Theorem 5.1. Moreover the sequence $\{X_n/n, n \in \mathbb{N}\}$ in Theorem 5.1 converges almost surely.*

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

In the application, $\mathcal{Z}_{J_X} \subset (-\infty, 1)$ corresponds to requiring the queue to be stable on the scale of large deviations. We also note that $x_\infty \in \mathcal{Z}_{J_X}$ if and only if

$$x_\infty = \sum_{k=1}^{K} v_k \phi_\infty^k \circ \psi_\infty^k \circ \left(\sum_{j=1}^{K} \phi_\infty^j \circ \psi_\infty^j \right)^{-1} \quad (1)$$

for some $(\phi_\infty^1, \ldots, \psi_\infty^K)$ such that $I(\phi_\infty^1, \ldots, \psi_\infty^K) = 0$, where I is the rate function in Assumption 4.1.

The following lemma is proved along the similar lines to results in the literature (see *e.g.* [17] and the references cited therein). However, here we consider the more general situation where $\mathcal{Z}_{J_X}$ is not reduced to a single point.

Lemma 6.1 (Tail asymptotics for supremum)**.** *Under Assumption 6.1, we have that*

$$-\inf_{s>1} J_{\sup X}(s) \leq \liminf_{n\to\infty} \frac{1}{n} \log P\left(X_{\sup} > n\right) \quad (6.1)$$

$$\leq \limsup_{n\to\infty} \frac{1}{n} \log P\left(X_{\sup} \geq n\right) \leq -\inf_{s\geq 1} J_{\sup X}(s), \quad (6.2)$$

where

$$J_{\sup X}(s) = \inf \left\{ I(\phi^1, \ldots, \psi^K) : \sup_{t>0} \left(\sum_{k=1}^{K} v_k \phi^k \circ \psi^k \circ \left(\sum_{j=1}^{K} \phi^j \circ \psi^j \right)^{-1}(t) - t \right) = s \right\}.$$

Note that the upper and lower bounds in equations (6.1) and (6.2) do not necessarily coincide. This is demonstrated by considering the following example of Benaim and Friz [5]: $P(X_{\sup} > x) = \exp(-\exp(\log[x]))$. Thus it is important to identify conditions such that $\inf_{s>1} J_{\sup X}(s) = \inf_{s\geq 1} J_{\sup X}(s)$. Due to the variational form of the rate function $J_{\sup X}$ we are able to give a positive answer if in place of Assumption 4.1 we consider an analogous assumption to Assumption 6.1 with either of Assumption 5.1 or Assumption 5.2. The values $v_1, \ldots, v_K$ that appear in the following assumption are those that occur in the definition of X_t presented in equation (5.1).

Assumption 6.2. *Assumption 5.1 holds with $I^1, \ldots, I^K : \mathbb{R}_+ \to [0, \infty)$, and $\sum_k v_k m_k l_k / \sum_j m_j l_j < 1$ for all $m_1, \ldots, m_K > 0$ and $(l_1, \ldots, l_K) \in \Delta_K$ such that $I^1(m_1) = \cdots = I^K(m_K) = H(l_1, \ldots, l_K) = 0$. Moreover the sequence $\{X_n/n, n \in \mathbb{N}\}$ in Theorem 5.1 converges almost surely.*

Proposition 6.2 (Tail asymptotics for supremum 1)**.** *Under Assumption 6.2, we have that* $\inf_{s>1} J_{\sup X}(s) = \inf_{s\geq 1} J_{\sup X}(s)$.

Assumption 6.3. *Assumption 5.2 holds with $I' : \mathbb{R}_+^K \to [0, \infty)$, and $\sum_k v_k m_k / \sum_j m_j < 1$ for all $(m_1, \ldots, m_K) \in \mathbb{R}_+^K$ such that $I'(m_1, \ldots, m_K) = 0$. Moreover the sequence $\{X_n/n, n \in \mathbb{N}\}$ in Theorem 5.1 converges almost surely.*

Proposition 6.3 (Tail asymptotics for supremum 2)**.** *Under Assumption 6.3, we have* $\inf_{s>1} J_{\sup X}(s) = \inf_{s\geq 1} J_{\sup X}(s)$.

Thus, under Assumptions 6.2 and 6.3 respectively, Propositions 6.2 and 6.3 prove that the two bounds in equations (6.1) and (6.2) coincide and that the tail of the distribution has logarithmic asymptotics:

$$\lim_{n\to\infty} \frac{1}{n} \log P(X_{\sup} > n) = -\inf_{s\geq 1} J_{\sup X}(s).$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

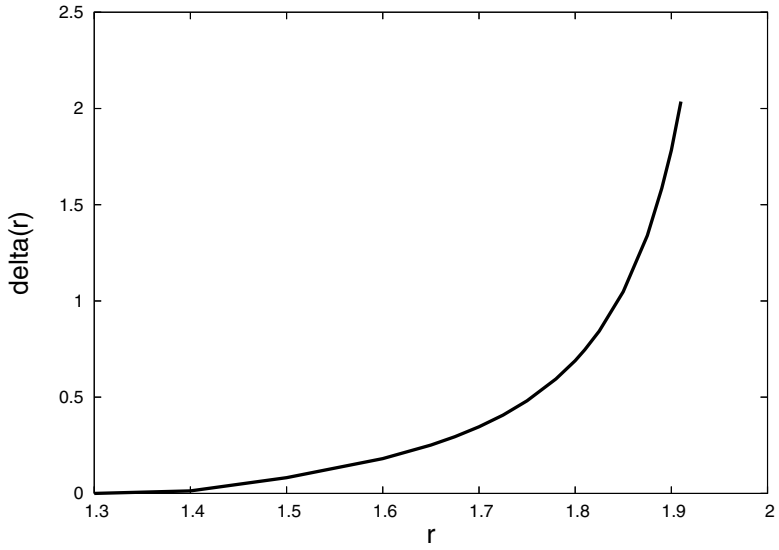


FIGURE 6. Rate of decay, $\delta(r)$, of the tail of the queuelength distribution at a single server queue fed with ITU-T P.59 and serving at rate r, where $\delta(r) = +\infty$ if $r > 2$

Example. Consider a single server queue processing at rate r and fed by a stochastic source of work, X_t, defined by the ITU-T P.59 process introduced in the example in Section 5.1. The queuelength distribution at the single server queue equals $\sup_{t>0} X_t - rt$ and, using the convexity of $J_X(\cdot)$ defined in equation (5.9), satisfies

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\sup_{t>0} X_t - rt > n\right) = -\inf_{s>0} sJ_X\left(\frac{1}{s} + r\right) =: -\delta(r).$$

That is, the tail of the queuelength distribution at a single server queue fed with ITU-T P.59 and served at rate r decays exponentially with rate $\delta(r)$ plotted in Figure 6. This can be used by teletraffic engineers as a quality of service metric for this source of traffic.

7. Modulated Lévy processes

In this section we consider independent Lévy processes selected by the phase process. This situation has interest in several fields. For instance it is used to model the price of a stock, where the phase process describes the underlying state of the economy.

Let $\{C_t^k,\, t \in \mathbb{R}_+\}_{k=1}^K$ be independent Lévy processes, independent of $\{Y_t,\, t \in \mathbb{R}_+\}$. We define the process $\{C_t,\, t \in \mathbb{R}_+\}$ as follows

$$C_t := \sum_{n=1}^{N_t} (C_{T_n}^{\sigma_n} - C_{T_{n-1}}^{\sigma_n}) + (C_t^{\sigma_{N_t+1}} - C_{T_{N_t}}^{\sigma_{N_t+1}}).$$

For any $k \in \{1, \ldots, K\}$ define the cumulant generating function $\Lambda_k(\gamma) := \log \mathbb{E}[\mathrm{e}^{\gamma C_1^k}]$ $(\gamma \in \mathbb{R})$ and its Legendre transform $\Lambda_k^*(c) := \sup_{\gamma\in\mathbb{R}}[\gamma c - \Lambda_k(\gamma)]$. Note that we can write the moment generating function of C_t in terms of Λ_{Y_s} as follows:

$$\mathbb{E}[\mathrm{e}^{\gamma C_t}] = \mathbb{E}[\mathrm{e}^{\sum_k \Lambda_k(\gamma) Z_t^k}] = \mathbb{E}[\mathrm{e}^{\int_0^t \Lambda_{Y_s}(\gamma)\mathrm{d}s}].$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Proposition 7.1 (LDP for modulated Lévy processes)**.** *If in addition to Assumption 4.1, the functions* $\Lambda_1, \ldots, \Lambda_K$ *are finite and essentially smooth, then*

(1) $\{(C_n/n, Z_n^1/n, \ldots, Z_n^K/n),\, n \in \mathbb{N}\}$ *satisfies the LDP in* $\mathbb{R}^{K+1}$ *with good rate function*

$$J_{C,\mathbf{Z}}(c, z_1, \ldots, z_K) = J_{C|\mathbf{Z}}(c|z_1, \ldots, z_K) + J_{\mathbf{Z}}(z_1, \ldots, z_K),$$

where $J_{C|\mathbf{Z}}(c|z_1, \ldots, z_K) = \inf\{\sum_{k=1}^K z_k \Lambda_k^*(c_k) : \sum_{k=1}^K z_k c_k = c\}$.

(2) $\{C_n/n,\, n \in \mathbb{N}\}$ *satisfies the LDP in* $\mathbb{R}$ *with good rate function*

$$J_C(c) = \inf\{J_{C|\mathbf{Z}}(c|z_1, \ldots, z_K) + J_{\mathbf{Z}}(z_1, \ldots, z_K) : (z_1, \ldots, z_K) \in \Delta_K\}.$$

Remark. The result in Proposition 7.1 generalizes a known result in the literature for discrete time Markov additive processes. Indeed, let $\{\sigma_n, n \in \mathbb{N}\}$ denote an irreducible Markov chain with state space $\{1, \ldots, K\}$ and transition matrix (p_{ij}), and assume that the sojourn times τ_n^k are all equal to 1. Thus $T_n = n$, $N_t = [t]$ and therefore

$$C_n = \sum_{k=1}^{n} (C_k^{\sigma_k} - C_{k-1}^{\sigma_k}).$$

Note that in this specific case $\{(\sigma_n, (C_n/n, Z_n^1/n, \ldots, Z_n^K/n)),\, n \in \mathbb{N}\}$ is a discrete time Markov additive process (with $K+1$ additive components). The rate function $J_{C,\mathbf{Z}}$ is given by a straightforward application of Proposition 7.1, where $J_{\mathbf{Z}}$ coincides with H on Δ_K and it is equal to infinity on Δ_K^c (see the remark at the end of Sect. 4.1). This is a well-known formula whose standard proof is based on a direct application of the Gärtner-Ellis theorem (see *e.g.* [22], [31], [32] and [25]).

8. Proofs

Lemma 3.1, $S \circ L$ representation.

Proof. We start by noting that, by (3.5), as $nt - [nt] < 1$ and $L_{[nt]}^k$ is an integer,

$$[nL_n^k(t)] = [L_{[nt]}^k + (nt - [nt])1_{\{k=k^*\}}] = L_{[nt]}^k.$$

Combining this with equations (3.6) and (3.5), we have

$$\begin{aligned} S_n^k(L_n^k(t)) &= \frac{1}{n} S^k_{[nL_n^k(t)]} + \left(L_n^k(t) - \frac{1}{n}[nL_n^k(t)]\right) \tau^k_{[nL_n^k(t)]+1} \\ &= \frac{1}{n} S^k_{L^k_{[nt]}} + \left(t - \frac{[nt]}{n}\right) 1_{\{k=k^*\}} \tau^k_{L^k_{[nt]}+1}. \end{aligned}$$

Finally, summing over all states $k \in \{1, \ldots, K\}$ gives

$$\sum_{k=1}^K S_n^k(L_n^k(t)) = \sum_{k=1}^K \frac{1}{n} S^k_{L^k_{[nt]}} + \left(t - \frac{[nt]}{n}\right) \tau^{k^*}_{L^{k^*}_{[nt]}+1} = T_n(t). \qquad \square$$

Lemma 3.2, $N_n = T_n^{-1}$.

Proof. We first show that $T_n(N_n(t)) = t$. Note that $[nN_n(t)] = N_{nt}$ as N_{nt} is an integer and

$$(nt - T_{N_{nt}}) \left(\tau^{k^{**}}_{L^{k^{**}}_{N_{nt}}+1}\right)^{-1} < 1.$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Therefore

$$k^*(n, N_n(t)) = \sigma_{[nN_n(t)]+1} = \sigma_{N_{nt}+1} = k^{**}(n,t) = k^{**},$$

where k^{**} is defined in equation (3.9). Using equations (3.7) and (3.10) we have that

$$\begin{aligned} T_n(N_n(t)) &= \frac{1}{n} T_{[nN_n(t)]} + \left(N_n(t) - \frac{[nN_n(t)]}{n} \right) \tau^{k^*(n,N_n(t))}_{L^{k^*(n,N_n(t))}_{[nN_n(t)]}+1} \\ &= \frac{1}{n} T_{N_{nt}} + \left(t - \frac{1}{n} T_{N_{nt}} \right) \left(\tau^{k^{**}}_{L^{k^{**}}_{N_{nt}}+1} \right)^{-1} \tau^{k^{**}}_{L^{k^{**}}_{N_{nt}}+1} = t. \end{aligned}$$

To show that the reverse property holds, $N_n(T_n(t)) = t$, first note that as $nt - [nt] < 1$

$$\begin{aligned} N_{nT_n(t)} &= \sup\{m : T_m \le nT_n(t)\} \\ &= \sup\left\{ m : T_m \le T_{[nt]} + (nt - [nt])\, \tau^{k^*}_{L^{k^*}_{[nt]}+1} \right\} \\ &= \sup\left\{ m : T_m \le T_{[nt]} \right\} = [nt] \end{aligned}$$

and that therefore

$$k^{**}(n, T_n(t)) = \sigma_{N_{nT_n(t)}+1} = \sigma_{[nt]+1} = k^*(n,t) = k^*,$$

where k^* is defined in (3.4). Using equations (3.7) and (3.10), this gives

$$\begin{aligned} N_n(T_n(t)) &= \frac{1}{n} N_{nT_n(t)} + \left(T_n(t) - \frac{T_{N_{nT_n(t)}}}{n} \right) \left(\tau^{k^{**}(n,T_n(t))}_{L^{k^{**}(n,T_n(t))}_{N_{nT_n(t)}}+1} \right)^{-1} \\ &= \frac{1}{n}[nt] + \left(T_n(t) - \frac{T_{[nt]}}{n} \right) \left(\tau^{k^*}_{L^{k^*}_{[nt]}+1} \right)^{-1} \\ &= \frac{1}{n}[nt] + \left(t - \frac{1}{n}[nt] \right) \tau^{k^*}_{L^{k^*}_{[nt]}+1} \left(\tau^{k^*}_{L^{k^*}_{[nt]}+1} \right)^{-1} = t. \end{aligned}$$

□

Theorem 3.3, $Z^k_n = S^k_n \circ L^k_n \circ N_n$.

Proof. As remarked in the proof of Lemma 3.2, we have $[nN_n(t)] = N_{nt}$ and $k^*(n, N_n(t)) = k^{**}$. Then, using equations (3.8) and (3.10), we obtain

$$\begin{aligned} S^k_n(L^k_n(N_n(t))) &= \frac{1}{n} S^k_{L^k_{[nN_n(t)]}} + \left(N_n(t) - \frac{1}{n}[nN_n(t)] \right) 1_{\{k=k^*(n,N_n(t))\}} \tau^k_{L^k_{[nN_n(t)]}+1} \\ &= \frac{1}{n} S^k_{L^k_{N_{nt}}} + \left(t - \frac{T_{N_{nt}}}{n} \right) \left(\tau^{k^{**}}_{L^{k^{**}}_{N_{nt}}+1} \right)^{-1} 1_{\{k=k^{**}\}} \tau^k_{L^k_{N_{nt}}+1} \\ &= \frac{1}{n} S^k_{L^k_{N_{nt}}} + \left(t - \frac{T_{N_{nt}}}{n} \right) 1_{\{k=k^{**}\}} = Z^k_n(t). \end{aligned}$$

Thus $Z^k_n(\cdot)$, defined in equation (3.11), is exactly $S^k_n \circ L^k_n \circ N_n(\cdot)$. We get the final result using the representation of $N_n(\cdot)$ in terms of $T_n(\cdot)$ in Lemma 3.2 and then the representation of $T_n(\cdot)$ in terms of S^k_n and L^k_n, $k \in \{1, \ldots, K\}$, in Lemma 3.1. □

Theorem 4.1, LDP for the empirical laws of the phase process.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Proof. We start proving that the function $\chi := \sum_{k=1}^{K} \phi^k \circ \psi^k$ belongs to $\mathcal{Y}_\uparrow$. The function χ is strictly increasing: for $t_1 < t_2$, we have $\psi^k(t_1) \leq \psi^k(t_2)$ and $\phi^k \circ \psi^k(t_1) \leq \phi^k \circ \psi^k(t_2)$ for all $k \in \{1, \ldots, K\}$; moreover, since $\sum_k \psi^k(t) = t$, there exists $\overline{k} \in \{1, \ldots, K\}$ such that $\psi^{\overline{k}}(t_1) < \psi^{\overline{k}}(t_2)$, and therefore $\phi^{\overline{k}} \circ \psi^{\overline{k}}(t_1) < \phi^{\overline{k}} \circ \psi^{\overline{k}}(t_2)$ since $\phi^{\overline{k}} \in \mathcal{Y}_\uparrow$. Furthermore there exists $\lim_{t\to\infty} \chi(t)(1+t)^{-1} \in (0, \infty)$; indeed, for each fixed $k \in \{1, \ldots, K\}$, there exist the limits $\psi^k(\infty) := \lim_t \psi^k(t) \in [0, \infty]$, $\ell_{\phi^k} := \lim_t \phi^k(t)(1+t)^{-1} \in (0, \infty)$ and $\ell_{\psi^k} := \lim_t \psi^k(t)(1+t)^{-1} \in [0, \infty)$ therefore we have $\lim_t \chi(t)(1+t)^{-1} = \sum_k \ell_{\phi^k} \ell_{\psi^k}$ since

$$\frac{\chi(t)}{1+t} = \sum_k \frac{\phi^k \circ \psi^k(t)}{1+\psi^k(t)} \frac{1+\psi^k(t)}{1+t}$$

and

$$\frac{\phi^k \circ \psi^k(t)}{1+\psi^k(t)} \frac{1+\psi^k(t)}{1+t} \to \begin{cases} \ell_{\phi^k} \ell_{\psi^k} & \text{if } \psi^k(\infty) = \infty \\ 0 & \text{if } \psi^k(\infty) < \infty \end{cases} \quad \text{as } t \to \infty;$$

moreover $\sum_k \ell_{\phi^k} \ell_{\psi^k} \in (0, \infty)$ since $\ell_{\psi^k} \in [0, \infty)$ and $\sum_k \ell_{\psi^k} = 1$.

Elementary arguments reveal that summation and projection are continuous on $\mathcal{Y}$. Inversion is continuous on $\mathcal{Y}_\uparrow$ by Lemma B.6 of Majewski [26]. Composition is continuous on $\mathcal{Y} \times \mathcal{Y}_\uparrow$ by Lemma B.1 of [26]. Thus, due to Theorem 3.3, we prove the statements 1 and 2 by applying the contraction principle with the following two maps in equations (8.1) and (8.2) respectively, which are continuous maps as they are the composition of continuous maps:

$$(\phi^1, \ldots, \phi^K, \psi^1, \ldots, \psi^K) \mapsto \left(\phi^1 \circ \psi^1 \circ \chi^{-1}, \ldots, \phi^K \circ \psi^K \circ \chi^{-1}\right) \tag{8.1}$$

$$\mapsto \left(\phi^1 \circ \psi^1 \circ \chi^{-1}, \ldots, \phi^K \circ \psi^K \circ \chi^{-1}\right)(1). \tag{8.2}$$

□

Theorem 5.1, LDP for fluid processes.

Proof. The LDP for $\{X_n/n,\, n \in \mathbb{N}\}$ follows from the representation in equation (5.3), statement 2 of Theorem 4.1 and an application of the contraction principle. □

Corollary 5.2, independent state selection and sojourn times.

Proof. We first prove the equality in equation (5.7). We start showing that

$$J_X(x) = \inf_{y>0} \inf \left\{ I(\phi^1, \ldots, \psi^K) : \sum_{k=1}^{K} v_k \phi^k(\psi^k(y)) = x, \sum_{k=1}^{K} \phi^k(\psi^k(y)) = 1 \right\}. \tag{8.3}$$

Define the sets

$$C_x := \left\{ (\phi^1, \ldots, \phi^K, \psi^1, \ldots, \psi^K) : \sum_{k=1}^{K} v_k\, \phi^k \circ \psi^k \circ \left(\sum_{j=1}^{K} \phi^j \circ \psi^j \right)^{-1} (1) = x \right\}$$

and

$$C_{x,y} := \left\{ (\phi^1, \ldots, \phi^K, \psi^1, \ldots, \psi^K) : \sum_{k=1}^{K} v_k \phi^k(\psi^k(y)) = x, \sum_{k=1}^{K} \phi^k(\psi^k(y)) = 1 \right\};$$

then $C_{x,y} \subset C_x$ for all $y > 0$, and therefore

$$J_X(x) \leq \inf_{y>0} \inf \left\{ I(\phi^1, \ldots, \psi^K) : (\phi^1, \ldots, \psi^K) \in C_{x,y} \right\}.$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Moreover the infimum $J_X(x)$ in (5.4) is attained at some $(\phi^1_*, \dots, \psi^K_*) \in C_x$ since C_x is closed and I is a good rate function; thus, defining $y^* = \left(\sum_k \phi^k_* \circ \psi^k_*\right)^{-1}(1)$, we have

$$\begin{aligned}\inf_{y>0} \inf \left\{ I(\phi^1, \dots, \psi^K) : (\phi^1, \dots, \psi^K) \in C_{x,y} \right\} &\le \inf \left\{ I(\phi^1, \dots, \psi^K) : (\phi^1, \dots, \psi^K) \in C_{x,y^*} \right\} \\ &\le I(\phi^1_*, \dots, \psi^K_*) = J_X(x).\end{aligned}$$

This proves (8.3). Arguing as for equality (8.3), we have

$$\begin{aligned}J_X(x) &= \inf_{y>0} \inf \left\{ I(\phi^1, \dots, \psi^K) : \sum_{k=1}^K v_k \phi^k(\psi^k(y)) = x, \sum_{k=1}^K \phi^k(\psi^k(y)) = 1 \right\} \\ &= \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf \left\{ I(\phi^1, \dots, \psi^K) : \right. \\ &\qquad \left. \psi^k(y) = y_k, \sum_{k=1}^K y_k = y, \sum_{k=1}^K v_k \phi^k(y_k) = x, \sum_{k=1}^K \phi^k(y_k) = 1 \right\} \\ &= \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf_{x_1,\dots,x_K>0} \inf \left\{ I(\phi^1, \dots, \psi^K) : \right. \\ &\qquad \left. \psi^k(y) = y_k, \sum_{k=1}^K y_k = y, \phi^k(y_k) = x_k, \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\} \\ &= \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf_{x_1,\dots,x_K>0} \left\{ \int_0^\infty \left(\sum_{k=1}^K I^k(\dot\phi^k(s)) + H(\dot\psi^1(s), \dots, \dot\psi^K(s)) \right) \mathrm{d}s : \right. \\ &\qquad \left. \psi^k(y) = y_k, \sum_{k=1}^K y_k = y, \phi^k(y_k) = x_k, \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\}.\end{aligned}$$

Therefore, defining

$$F(y, y_1, \dots, y_K, x_1, \dots, x_K) := \sum_{k=1}^K y_k I^k\left(\frac{x_k}{y_k}\right) + yH\left(\frac{y_1}{y}, \dots, \frac{y_K}{y}\right)$$

for $y, y_1, \dots, y_K, x_1, \dots, x_K > 0$, we have

$$\begin{aligned}J_X(x) &\ge \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf_{x_1,\dots,x_K>0} \left\{ \sum_{k=1}^K \int_0^{y_k} I^k(\dot\phi^k(s))\mathrm{d}s + \int_0^y H(\dot\psi^1(s), \dots, \dot\psi^K(s))\mathrm{d}s : \right. \\ &\qquad \left. \psi^k(y) = y_k, \sum_{k=1}^K y_k = y, \phi^k(y_k) = x_k, \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\} \\ &\ge \inf_{y>0} \inf_{y_1,\dots,y_K>0} \inf_{x_1,\dots,x_K>0} \{ F(y, y_1, \dots, y_K, x_1, \dots, x_K) : \\ &\qquad \left. \sum_{k=1}^K y_k = y, \sum_{k=1}^K v_k x_k = x, \sum_{k=1}^K x_k = 1 \right\} =: \widetilde{J}(x), \end{aligned} \tag{8.4}$$

where in the last line we have used Jensen's inequality, as I^k and H are convex.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

By the definition of the infimum $\widetilde{J}(x)$ in (8.4), given $\varepsilon > 0$, there exist $\hat{y}, \hat{y}_1, \ldots, \hat{y}_K, \hat{x}_1, \ldots \hat{x}_K > 0$ such that $\sum \hat{y}_k = \hat{y}$, $\sum v_k \hat{x}_k = x$ and $\sum \hat{x}_k = 1$ and

$$F(\hat{y}, \hat{y}_1, \ldots, \hat{y}_K, \hat{x}_1, \ldots, \hat{x}_K) \leq \widetilde{J}(x) + \varepsilon.$$

Note that I attains the value 0 since it is a good rate function; thus there exist $m_1, \ldots, m_K \in (0, \infty)$ and $(l_1, \ldots, l_K) \in \Delta_K$ such that $I^k(m_k) = 0$ ($k \in \{1, \ldots, K\}$) and $H(l_1, \ldots, l_K) = 0$. For $k \in \{1, \ldots, K\}$ we define the functions

$$\hat{\phi}^k(t) = \begin{cases} \dfrac{\hat{x}_k}{\hat{y}_k} t & \text{if } t \leq \hat{y}_k \\ \hat{x}_k + (t - \hat{y}_k) m_k & \text{if } t \geq \hat{y}_k \end{cases} \quad \text{and} \quad \hat{\psi}^k(t) = \begin{cases} \dfrac{\hat{y}_k}{\hat{y}} t & \text{if } t \leq \hat{y} \\ \hat{y}_k + (t - \hat{y}) l_k & \text{if } t \geq \hat{y}; \end{cases}$$

then it is easy to check that

$$\hat{\phi}^k(\hat{\psi}^k(t)) = \begin{cases} \dfrac{\hat{x}_k}{\hat{y}} t & \text{if } t \leq \hat{y} \\ \hat{x}_k + (t - \hat{y}) l_k m_k & \text{if } t \geq \hat{y}, \end{cases}$$

these functions meet the constraints of equation (5.4) and, using equation (5.6), we have that

$$J_X(x) \leq I(\hat{\phi}^1, \ldots, \hat{\phi}^K, \hat{\psi}^1, \ldots, \hat{\psi}^K) = F(\hat{y}, \hat{y}_1, \ldots, \hat{y}_K, \hat{x}_1, \ldots, \hat{x}_K) \leq \widetilde{J}(x) + \varepsilon.$$

Thus $J_X(x) = \widetilde{J}(x)$ since $\varepsilon > 0$ is arbitrary, and this proves (5.7).

We finally show that J_X is convex. We start showing that F defined above is convex. For $\alpha \in [0, 1]$ consider

$$F(\alpha y + (1-\alpha)u, \alpha y_1 + (1-\alpha)u_1, \ldots, \alpha y_K + (1-\alpha)u_K, \alpha x_1 + (1-\alpha)w_1, \ldots, \alpha x_K + (1-\alpha)w_K).$$

This equals

$$\begin{aligned} &\sum_{k=1}^{K} (\alpha y_k + (1-\alpha)u_k) I^k \left(\frac{\alpha x_k + (1-\alpha)w_k}{\alpha y_k + (1-\alpha)u_k} \right) \\ &\qquad + (\alpha y + (1-\alpha)u) H \left(\frac{\alpha y_1 + (1-\alpha)u_1}{\alpha y + (1-\alpha)u}, \ldots, \frac{\alpha y_K + (1-\alpha)u_K}{\alpha y + (1-\alpha)u} \right). \end{aligned} \tag{8.5}$$

Now, defining

$$\gamma = \frac{\alpha y}{\alpha y + (1-\alpha)u} \in [0, 1] \text{ and for all } k \in \{1, \ldots, K\}\ \gamma_k = \frac{\alpha y_k}{\alpha y_k + (1-\alpha)u_k} \in [0, 1],$$

the expression (8.5) equals

$$\begin{aligned} &\sum_{k=1}^{K} (\alpha y_k + (1-\alpha)u_k) I^k \left(\gamma_k \frac{x_k}{y_k} + (1-\gamma_k) \frac{w_k}{u_k} \right) \\ &\qquad + (\alpha y + (1-\alpha)u) H \left(\gamma \frac{y_1}{y} + (1-\gamma) \frac{u_1}{u}, \ldots, \gamma \frac{y_K}{y} + (1-\gamma) \frac{u_K}{u} \right). \end{aligned}$$

The convexity of $I^1, \ldots, I^K$ and H easily gives the convexity of F.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Now, define the $(2k+1)\times 3$ matrix

$$A = \begin{pmatrix} -1 & 0 & 0 \\ +1 & 0 & 0 \\ & \vdots & \\ +1 & 0 & 0 \\ 0 & v_1 & 1 \\ & \vdots & \\ 0 & v_K & 1 \end{pmatrix}$$

and the function

$$\widetilde{F}(w,x,z) = \inf\{F(y,y_1,\dots,y_K,x_1,\dots,x_K) : (y,y_1,\dots,y_K,x_1,\dots,x_K)A = (w,x,z)\}$$

where the infimum is over all positive $y, y_1,\dots,y_K,x_1,\dots,x_K$; then $\widetilde{F}$ is convex by the convexity of F and Theorem 5.7 of Rockafellar [38]. Therefore $J_X(x) = \widetilde{F}(0,x,1)$ is also convex. □

Corollary 5.3, equally visited states.

Proof. The proof of this corollary is similar to that for Corollary 5.2 and is omitted. □

Lemma 6.1, tail asymptotics for supremum.

Proof. We shall show later that the sequence $\{\sup_{t>0}\{X_n(t)-t\}, n\in\mathbb{N}\}$ satisfies the LDP with good rate function $J_{\sup X}$. Thus the claim is an immediate consequence of this LDP and the equalities

$$\{X_{\sup} > n\} = \left\{\sup_{t>0}(X_n(t)-t) > 1\right\} \text{ and } \{X_{\sup}\geq n\} = \left\{\sup_{t>0}(X_n(t)-t)\geq 1\right\},$$

which hold as $X_{nt}/n = X_n(t)$ for all $n\in\mathbb{N}$ and $t>0$. It remains to prove the announced LDP. Note that

$$P\left(\left\{\lim_{n\to\infty}\frac{X_n}{n}\in\mathcal{Z}_{J_X}\right\}\right) = 1 \tag{8.6}$$

by Theorem 5.1 and Lemma A.1 (for Lem. A.1, see the Appendix at the end of this section). Let $\mathcal{Z}_{J_X}-1$ be the set defined by

$$\mathcal{Z}_{J_X}-1 := \{x\in\mathbb{R} : x = y-1, y\in\mathcal{Z}_{J_X}\}.$$

The set $\mathcal{Z}_{J_X}-1$ is compact since $\mathcal{Z}_{J_X}$ is the set of all the zeros of the good rate function J_X; thus

$$\widetilde{\mathcal{Y}} := \left\{\phi\in\mathcal{Y} : \lim_{t\to\infty}\frac{\phi(t)}{1+t}\in\mathcal{Z}_{J_X}-1\right\}$$

is closed. We remark that, for each $n\in\mathbb{N}$, $X_n(\cdot)-\mathrm{id}(\cdot)$ (where $\mathrm{id}(t)=t$) belongs to $\widetilde{\mathcal{Y}}$ almost surely; indeed $X_n(\cdot)-\mathrm{id}(\cdot)$ is continuous, $X_n(\cdot)-\mathrm{id}(\cdot)$ starts at the origin (*i.e.* $X_n(0)-\mathrm{id}(0)=0$) and, by (8.6),

$$\lim_{t\to\infty}\frac{X_n(t)-\mathrm{id}(t)}{1+t} = \lim_{t\to\infty}\frac{X_{nt}}{n(1+t)} - \frac{t}{1+t}\in\mathcal{Z}_{J_X}-1$$

almost surely.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Arguing as in the proof of Theorem 4.1 one can show that the function $(\phi^1,\dots,\psi^K) \mapsto \sum_k \phi^k \circ \psi^k \circ \left(\sum_j \phi^j \circ \psi^j\right)^{-1}$ is continuous. Thus, by Assumption 4.1 and the contraction principle, $\{X_n(\cdot) - \mathrm{id}(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $\widetilde{\mathcal{Y}}$ with good rate function $J_{X(\bullet)-\mathrm{id}}$ defined by

$$J_{X(\bullet)-\mathrm{id}}(\eta) = \inf\left\{ I(\phi^1,\dots,\psi^K) : \sum_{k=1}^K v_k\, \phi^k \circ \psi^k \circ \left(\sum_{j=1}^K \phi^j \circ \psi^j\right)^{-1} - \mathrm{id} = \eta \right\}.$$

Finally the LDP follows by applying again the contraction principle. Indeed, the function $\eta \mapsto \sup\{\eta(t) : t > 0\}$ is continuous on $\widetilde{\mathcal{Y}}$ since $\mathcal{Z}_{J_X} - 1 \subset (-\infty, 0)$ by Assumption 6.1; this can be checked by adapting the proof of Theorem 5.3 in [17]. □

Proposition 6.2, tail asymptotics for supremum 1.

Proof. First note that the hypotheses imply Assumption 6.1. We distinguish two cases.
Case 1: $v_k \le 1$ for all $k \in \{1,\dots,K\}$. We have $P\left(\sup_{t>0}\{X_t - t\} = 0\right) = 1$, and therefore

$$J_{\sup X}(s) = \begin{cases} 0 & \text{if } s = 0 \\ \infty & \text{if } s \neq 0; \end{cases}$$

thus $\inf_{s>1} J_{\sup X}(s) = \inf_{s\ge 1} J_{\sup X}(s) = \infty$.
Case 2: there exists $k^* \in \{1,\dots,K\}$ such that $v_{k^*} > 1$. In this case $J_{\sup X}$ has the following form:

$$J_{\sup X}(s) = \inf\left\{ \int_0^\infty \left(\sum_{k=1}^K I^k(\dot\phi^k(t)) + H(\dot\psi^1(t),\dots,\dot\psi^K(t))\right) \mathrm{d}t : \right.$$
$$\left. (\phi^1,\dots,\psi^K) \in \mathcal{Y}^{\#}, \sup_{t>0}\left(\sum_{k=1}^K v_k \phi^k \circ \psi^k \circ \left(\sum_{j=1}^K \phi^j \circ \psi^j\right)^{-1}(t) - t\right) = s \right\}.$$

We remark that, since the infimum over a closed set of a good rate function is attained at some point, we have: $\inf_{s\ge 1} J_{\sup X}(s) = J_{\sup X}(s^*)$ for some $s^* \ge 1$;

$$J_{\sup X}(s^*) = \int_0^\infty \left(\sum_{k=1}^K I^k(\dot\phi^k_*(t)) + H(\dot\psi^1_*(t),\dots,\dot\psi^K_*(t))\right) \mathrm{d}t$$

for some $(\phi^1_*,\dots,\psi^K_*) \in \mathcal{Y}^{\#}$ such that

$$\sup_{t>0}\left(\sum_{k=1}^K v_k \phi^k_* \circ \psi^k_* \circ \left(\sum_{j=1}^K \phi^j_* \circ \psi^j_*\right)^{-1}(t) - t\right) = s^*.$$

We remark that, for $(\phi^1,\dots,\psi^K) \in \mathcal{Y}^{\#}$ and $y > 0$, by Jensen's inequality we have

$$\int_0^\infty \left(\sum_{k=1}^K I^k(\dot\phi^k(s)) + H(\dot\psi^1(s),\dots,\dot\psi^K(s))\right) \mathrm{d}s \ge \sum_{k=1}^K y_k I^k\left(\frac{x_k}{y_k}\right) + yH\left(\frac{y_1}{y},\dots,\frac{y_k}{y}\right)$$

where $\psi^k(y) = y_k$ and $x_k = \phi^k(y_k)$ for all $k \in \{1,\dots,K\}$; moreover the lower bound is attained with suitable piecewise linear functions. We recall that there exist $m_1,\dots,m_K > 0$ and $(l_1,\dots,l_K) \in \Delta_K$ such that $I^1(m_1) =$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

$\cdots = I^K(m_K) = H(l_1, \ldots, l_K) = 0$; for a choice of these values we set $x_\infty = \sum_k v_k m_k l_k / \sum_j m_j l_j$, and we have $x_\infty < 1$ by Assumption 6.2. Thus the minimizing point $(\phi_*^1, \ldots, \psi_*^K) \in \mathcal{Y}^\#$ is of the form

$$\phi_*^k(t) = \begin{cases} \dfrac{x_k}{y_k} t & \text{if } t \le y_k \\ x_k + (t - y_k) m_k & \text{if } t \ge y_k \end{cases} \quad \text{and } \psi_*^k(t) = \begin{cases} \dfrac{y_k}{y} t & \text{if } t \le y \\ y_k + (t - y) l_k & \text{if } t \ge y \end{cases}$$

for $k \in \{1, \ldots, K\}$ and $y, y_1, \ldots, y_K, x_1, \ldots, x_K > 0$ such that $\sum_k y_k = y$. Moreover we have to choose $x_1, \ldots, x_K > 0$ such that $\sum_k (v_k - 1) x_k = s^*$; indeed it is easy to check that

$$\phi_*^k \circ \psi_*^k(t) = \begin{cases} \dfrac{x_k}{y} t & \text{if } t \le y \\ x_k + (t - y) l_k m_k & \text{if } t \ge y \end{cases} \quad (k \in \{1, \ldots, K\}),$$

$$\left(\sum_j \phi_*^j \circ \psi_*^j \right)^{-1}(t) = \begin{cases} \dfrac{y}{\sum_j x_j} t & \text{if } t \le \sum_j x_j \\ y + \dfrac{t - \sum_j x_j}{\sum_j m_j l_j} & \text{if } t \ge \sum_j x_j, \end{cases}$$

$$\sum_k v_k \phi_*^k \circ \psi_*^k \circ \left(\sum_j \phi_*^j \circ \psi_*^j \right)^{-1}(t) = \begin{cases} \dfrac{\sum_k v_k x_k}{\sum_j x_j} t & \text{if } t \le \sum_j x_j \\ \sum_k v_k \left(x_k + \dfrac{t - \sum_j x_j}{\sum_j m_j l_j} m_k l_k \right) & \text{if } t \ge \sum_j x_j \end{cases}$$

$$= \begin{cases} \left(\dfrac{s^*}{\sum_j x_j} + 1 \right) t & \text{if } t \le \sum_j x_j \\ s^* + \sum_j x_j + x_\infty (t - \sum_j x_j) & \text{if } t \ge \sum_j x_j, \end{cases}$$

and therefore

$$\sup_{t > 0} \left(\sum_k v_k \phi_*^k \circ \psi_*^k \circ \left(\sum_j \phi_*^j \circ \psi_*^j \right)^{-1}(t) - t \right) = s^*$$

since $x_\infty < 1$.

Now, for a fixed $\varepsilon > 0$, consider $(\phi_\varepsilon^1, \ldots, \phi_\varepsilon^K, \psi_\varepsilon^1, \ldots, \psi_\varepsilon^K)$ such that $\psi_\varepsilon^k = \psi_*^k$ for all $k \in \{1, \ldots, K\}$, $\phi_\varepsilon^k = \phi_*^k$ for all $k \in \{1, \ldots, K\} \setminus \{k^*\}$ and

$$\phi_\varepsilon^{k^*}(t) = \begin{cases} \dfrac{x_{k^*} + \varepsilon}{y_{k^*}} t & \text{if } t \le y_{k^*} \\ x_{k^*} + \varepsilon + (t - y_{k^*}) m_{k^*} & \text{if } t \ge y_{k^*}. \end{cases}$$

It is easy to check that

$$\sum_k v_k \phi_\varepsilon^k \circ \psi_\varepsilon^k \circ \left(\sum_j \phi_\varepsilon^j \circ \psi_\varepsilon^j \right)^{-1}(t) - t = \begin{cases} \dfrac{s^* + (v_{k^*} - 1)\varepsilon}{\sum_j x_j + \varepsilon} t & \text{if } t \le \sum_j x_j + \varepsilon \\ s^* + (v_{k^*} - 1)\varepsilon & \\ + (x_\infty - 1) \left(t - \left(\sum_j x_j + \varepsilon \right) \right) & \text{if } t \ge \sum_j x_j + \varepsilon, \end{cases}$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

and therefore

$$\sup_{t>0}\left(\sum_k v_k\phi_\varepsilon^k\circ\psi_\varepsilon^k\circ\left(\sum_j\phi_\varepsilon^j\circ\psi_\varepsilon^j\right)^{-1}(t)-t\right)=s^*+(v_{k^*}-1)\varepsilon$$
$$>s^*\geq 1,$$

since $x_\infty<1$, $v_{k^*}>1$ and $\varepsilon>0$. In conclusion we have

$$\begin{aligned}\inf_{s\geq 1}J_{\sup X}(s)&\leq\inf_{s>1}J_{\sup X}(s)\\&\leq I(\phi_\varepsilon^1,\ldots,\phi_\varepsilon^K,\psi_\varepsilon^1,\ldots,\psi_\varepsilon^K)\\&=\sum_{k\neq k^*}y_kI^k\left(\frac{x_k}{y_k}\right)+y_{k^*}I^{k^*}\left(\frac{x_{k^*}+\varepsilon}{y_{k^*}}\right)+yH\left(\frac{y_1}{y},\ldots,\frac{y_k}{y}\right)\\&\to_{\varepsilon\to 0}\sum_{k=1}^K y_kI^k\left(\frac{x_k}{y_k}\right)+yH\left(\frac{y_1}{y},\ldots,\frac{y_k}{y}\right)\qquad(8.7)\\&=I(\phi_*^1,\ldots,\psi_*^K)=J_{\sup X}(s^*)=\inf_{s\geq 1}J_{\sup X}(s),\end{aligned}$$

where in (8.7) we used the continuity of I^{k^*}. □

Proposition 6.3, tail asymptotics for supremum 2.

Proof. The proof of this proposition is similar to that for Proposition 6.2 and is omitted. □

Proposition 7.1, LDP for modulated Lévy processes.

Proof. We only need to show that statement 1 holds, as statement 2 is a consequence of the statement 1 and the contraction principle. The proof of the statement 1 is based on Theorem 2.3 of Chaganty [9]. Denote by $\nu_n(\cdot|Z_n^1/n,\ldots,Z_n^K/n)$ the conditional law of C_n/n given $(Z_n^1/n,\ldots,Z_n^K/n)$. The sequence $\{(Z_n^1/n,\ldots,Z_n^K/n), n\in\mathbb{N}\}$ satisfies the LDP with good rate function $J_{\mathbf{Z}}$ by statement 2 of Theorem 4.1. Thus we only have to check the following three conditions: (i) for any sequence $\{\mathbf{z}_n\}\subset\Delta_K$ which converges to a point $\mathbf{z}=(z^1,\ldots,z^K)\in\Delta_K$, $\{\nu_n(\cdot|\mathbf{z}_n)\}$ satisfies the LDP with good rate function $J_{C|\mathbf{Z}}(\cdot|z^1,\ldots,z^K)$; (ii) the function $(c,z^1,\ldots,z^K)\mapsto J_{C|\mathbf{Z}}(c|z^1,\ldots,z^K)$ is lower semi-continuous; (iii) $J_{C,\mathbf{Z}}$ is a good rate function.

Proof of (i). Denote by $\widehat{\nu}_n(\cdot|\mathbf{z}_n)$ the moment generating function concerning $\nu_n(\cdot|\mathbf{z}_n)$. Then

$$\lim_{n\to\infty}\frac{1}{n}\log\widehat{\nu}_n(n\gamma|\mathbf{z}_n)=\lim_{n\to\infty}\frac{1}{n}\log\exp\left(n\sum_k\Lambda_k(\gamma)z_n^k\right)=\sum_k z^k\Lambda_k(\gamma).$$

So, by Gärtner Ellis Theorem (see *e.g.* [12], Chap. 2, Sect. 3), $\{\nu_n(\cdot|\mathbf{z}_n)\}$ satisfies the LDP with good rate function

$$J_{C|\mathbf{Z}}(c|z^1,\ldots,z^K)=\sup_{\gamma\in\mathbb{R}}\left[\gamma c-\sum_k z^k\Lambda_k(\gamma)\right].$$

We complete the proof of *(i)* by noting that the Legendre transform of a sum of lower semi-continuous convex functions is the infimal convolution of the Legendre transforms of the individual elements (*e.g.* Thm. 16.4

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

of [38]) and that $\sup_{\gamma\in\mathbb{R}}\{\gamma x - z^k\Lambda_k(\gamma)\} = z^k\Lambda^*(x/z^k)$. Therefore we have that

$$J_{C|\mathbf{Z}}(c|z^1,\dots,z^K) = \inf\left\{\sum_{k=1}^K z_k\Lambda^*(c_k) : \sum_{k=1}^K z_k c_k = c\right\}.$$

Proof of (ii). Consider a sequence $\{(c_n,\mathbf{z}_n)\}\subset\mathbb{R}\times\Delta_K$ which converges to a point $(c,\mathbf{z})\in\mathbb{R}\times\Delta_K$. Then, for all $\gamma\in\mathbb{R}$, we have $J_{C|\mathbf{Z}}(c_n|z_n^1,\dots,z_n^K)\geq\gamma c_n - \sum_k z_n^k\Lambda_k(\gamma)$ (for all $n\in\mathbb{N}$), and therefore

$$\liminf_{n\to\infty} J_{C|\mathbf{Z}}(c_n|z_n^1,\dots,z_n^K)\geq\gamma c - \sum_k z^k\Lambda_k(\gamma).$$

The claim follows taking the supremum over $\gamma\in\mathbb{R}$.
Proof of (iii). We have to check that, for all $\eta>0$, the level set $\mathcal{L}_\eta := \{(c,\mathbf{z}) : J_{C,\mathbf{Z}}(c,\mathbf{z})\leq\eta\}$ is compact. First of all it is a closed set since $J_{C,\mathbf{Z}}$ is lower semi-continuous; indeed, for any sequence $\{(c_n,\mathbf{z}_n)\}\subset\mathbb{R}\times\Delta_K$ which converges to a point $(c,\mathbf{z})\in\mathbb{R}\times\Delta_K$, the lower semi-continuity of $J_{C|\mathbf{Z}}$ and $J_{\mathbf{Z}}$ gives

$$\begin{aligned}\liminf_{n\to\infty} J_{C,\mathbf{Z}}(c_n,\mathbf{z}_n) &\geq \liminf_{n\to\infty} J_{C|\mathbf{Z}}(c_n|\mathbf{z}_n) + \liminf_{n\to\infty} J_{\mathbf{Z}}(\mathbf{z}_n)\\ &\geq J_{C|\mathbf{Z}}(c|\mathbf{z}) + J_{\mathbf{Z}}(\mathbf{z}) = J_{C,\mathbf{Z}}(c,\mathbf{z}).\end{aligned}$$

Now consider the set $G_\eta := \{(c,\mathbf{z})\in\mathbb{R}\times\Delta_K : J_{C|\mathbf{Z}}(c|\mathbf{z})\leq\eta\}$ and note that $\mathcal{L}_\eta\subset G_\eta$. The claim follows if we prove that G_η is bounded. Reasoning by contradiction, since Δ_K is bounded, there exists a sequence $\{(c_n,\mathbf{z}_n)\}\subset G_\eta$ such that $\lim_{n\to\infty}|c_n| = +\infty$. Thus we have

$$\eta\geq J_{C|\mathbf{Z}}(c_n|\mathbf{z}_n)\geq\gamma c_n - \sum_k z_n^k\Lambda_k(\gamma)$$

for all $\gamma\in\mathbb{R}$ and for all $n\in\mathbb{N}$; this leads to a contradiction if we choose $\gamma\in\mathbb{R}$ such that $\lim_{n\to\infty}\gamma c_n = +\infty$. □

Appendix A. Lemma A.1

Lemma A.1. *Let $\{W_n,\, n\in\mathbb{N}\}$ be a sequence of real valued random variables which satisfies the LDP with good rate function I, and let $\mathcal{Z}$ be the set defined by $\mathcal{Z} := \{w\in\mathbb{R} : I(w) = 0\}$. Moreover assume that $P(\{\lim_{n\to\infty} W_n = W_\infty\}) = 1$ for a suitable random variable W_∞. Then $P(W_\infty\in\mathcal{Z}) = 1$.*

Proof. In this proof we use the following standard notation for $A\subset\mathbb{R}$: $A^c = \mathbb{R}\setminus A$ and $\overline{A}$ is the closure of A. The set $\mathcal{Z}$ is closed and nonempty since I is a good rate function. Let $\mathcal{Z}_\varepsilon$ be the set defined by $\mathcal{Z}_\varepsilon := \{x\in\mathbb{R} : \text{there exists } y_x\in\mathcal{Z} \text{ such that } |x-y_x|<\varepsilon\}$ for $\varepsilon>0$; then $I(\overline{\mathcal{Z}_\varepsilon^c}) := \inf\{I(x) : x\in\overline{\mathcal{Z}_\varepsilon^c}\} > 0$.

We have $\limsup_{n\to\infty}\frac{1}{n}\log P(W_n\in\overline{\mathcal{Z}_\varepsilon^c})\leq -I(\overline{\mathcal{Z}_\varepsilon^c})$ and, for $\delta\in(0, I(\overline{\mathcal{Z}_\varepsilon^c}))$, there exists $\overline{n}\in\mathbb{N}$ such that $P(W_n\in\overline{\mathcal{Z}_\varepsilon^c})\leq \mathrm{e}^{-n(I(\overline{\mathcal{Z}_\varepsilon^c})-\delta)}$ for $n\geq\overline{n}$. Then $\sum_{n\geq\overline{n}} P(W_n\in\overline{\mathcal{Z}_\varepsilon^c})\leq\sum_{n\geq\overline{n}}\mathrm{e}^{-n(I(\overline{\mathcal{Z}_\varepsilon^c})-\delta)}<\infty$ and, by the Borel-Cantelli lemma, we have $P\left(\liminf_{n\to\infty}\left\{W_n\notin\overline{\mathcal{Z}_\varepsilon^c}\right\}\right) = 1$. Thus $P\left(\liminf_{n\to\infty}\{W_n\in\mathcal{Z}_\varepsilon\}\right) = 1$ since $\overline{\mathcal{Z}_\varepsilon^c}\supset\mathcal{Z}_\varepsilon^c$. Then we have $P(A_\varepsilon) = 1$ where

$$\begin{aligned}A_\varepsilon &= \liminf_{n\to\infty}\{|W_n - W_\infty|<\varepsilon\}\cap\liminf_{n\to\infty}\{W_n\in\mathcal{Z}_\varepsilon\}\\ &= \liminf_{n\to\infty}\left(\{|W_n - W_\infty|<\varepsilon\}\cap\{W_n\in\mathcal{Z}_\varepsilon\}\right);\end{aligned}$$

this means that there exists $n_\varepsilon^*(\omega)$ such that for all $n\geq n_\varepsilon^*(\omega)$ we have $|W_n(\omega) - W_\infty(\omega)|<\varepsilon$ and $|W_n(\omega) - y_{W_n(\omega)}|<\varepsilon$ for some $y_{W_n(\omega)}\in\mathcal{Z}$, whence we obtain

$$|W_\infty(\omega) - y_{W_n(\omega)}|\leq|W_\infty(\omega) - W_n(\omega)| + |W_n(\omega) - y_{W_n(\omega)}|<2\varepsilon.$$

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

Therefore $W_\infty(\omega) \in \mathcal{Z}_{2\varepsilon}$ (with $y_{W_\infty(\omega)} = y_{W_n(\omega)} \in \mathcal{Z}$ for some arbitrarily fixed $n \geq n^*_\varepsilon(\omega)$); thus $A_\varepsilon \subset \{W_\infty \in \mathcal{Z}_{2\varepsilon}\}$, and therefore $P(W_\infty \in \mathcal{Z}_{2\varepsilon}) = 1$. We conclude the proof noting that $P(W_\infty \in \mathcal{Z}) = 1$. Indeed, since $\mathcal{Z}$ is a closed set, we have $\mathcal{Z} = \bigcap_{\varepsilon>0,\varepsilon\in\mathbb{Q}} \mathcal{Z}_{2\varepsilon}$. □

Acknowledgements. The authors thank an anonymous reviewer for spotting an error in, and proposing the correction to, our original definition of $J_{C|\mathbf{Z}}(c|z_1,\ldots,z_K)$ in Proposition 7.1.

References

[1] S. Asmussen, Risk theory in a Markovian environment. *Scand. Actuar. J.* (1989) 69–100.
[2] S. Asmussen, *Ruin probabilities*. Advanced Series on Statistical Science & Applied Probability, Vol. 2, World Scientific Publishing Co. Inc., River Edge, NJ (2000).
[3] S. Asmussen, *Applied probability and queues*, second edn., Applications of Mathematics (New York), Vol. 51, Springer-Verlag, New York (2003), Stochastic Modelling and Applied Probability.
[4] S. Asmussen and M. Pihlsgård, Loss rates for Lévy processes with two reflecting barriers. *Math. Oper. Res.* **32** (2007) 308–321.
[5] S. Benaim and P. Friz, Smile asymptotics. II. Models with known moment generating functions. *J. Appl. Probab.* **45** (2008) 16–32.
[6] P. Billingsley, *Convergence of probability measures*, Wiley Inter-Science (1999).
[7] P.T. Brady, A statistical analysis of on-off patterns in 16 conversations. *The Bell Systems Technical Journal* **47** (1968) 73–91.
[8] L. Breuer, On Markov-additive jump processes. *Queueing Syst.* **40** (2002) 75–91.
[9] N.R. Chaganty, Large deviations for joint distributions and statistical applications. *Sankhyā Ser.* A **59** (1997) 147–166.
[10] E. Çinlar, Markov additive processes. I, II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **24** (1972) 85–93; E. Çinlar, Markov additive processes. I, II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** (1972) 95–121.
[11] A. Dembo and T. Zajic, Large deviations: from empirical mean and measure to partial sums process. *Stochastic Process. Appl.* **57** (1995) 191–224.
[12] A. Dembo and O. Zeitouni, *Large deviation techniques and applications*. Springer (1998).
[13] J-D. Deuschel and D.W. Stroock, *Large deviations*. Academic Press (1989).
[14] J.H. Dshalalow, Characterization of modulated Cox measures on topological spaces. *Int. J. Appl. Math. Stat.* **11** (2007) 21–37.
[15] J.H. Dshalalow and G. Russell, *On a single-server queue with fixed accumulation level, state dependent service, and semi-Markov modulated input flow. Internat. J. Math. Math. Sci.* **15** (1992) 593–600.
[16] K.R. Duffy and A. Sapozhnikov, The large deviation principle for the on-off Weibull sojourn process. *J. Appl. Probab.* **45** (2008) 107–117.
[17] A. Ganesh, N. O'Connell and D. Wischik, *Big queues*, Lecture Notes in Mathematics, Vol. 1838. Springer-Verlag, Berlin (2004).
[18] A.J. Ganesh and N. O'Connell, A large deviation principle with queueing applications. *Stochastics and Stochastic Reports* **73** (2002) 25–35.
[19] N. Gantert, Functional Erdős-Renyi laws for semiexponential random variables. *Ann. Probab.* **26** (1998) 1356–1369.
[20] J. Garcia, An extension of the contraction principle. *J. Theoret. Probab.* **17** (2004) 403–434.
[21] H. Heffes and D.M. Luncantoni, A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE Journal on Selected Areas in Communications* **4** (1986) 856–868.
[22] I. Iscoe, P. Ney and E. Nummelin, Large deviations of uniformly recurrent Markov additive processes. *Adv. in Appl. Math.* **6** (1985) 373–412.
[23] G. Latouche, M. Remiche and P. Taylor, Transient Markov arrival processes. *Ann. Appl. Probab.* **13** (2003) 628–640.
[24] H.H. Lee and C.K. Un, A study of on-off characteristics of conversation speech. *IEEE Transactions on Communications* **34** (1986) 630–636.
[25] T. Lehtonen and H. Nyrhinen, On asymptotically efficient simulation of ruin probabilities in a Markovian environment. *Scand. Actuar. J.* (1992) 60–75. MR MR1193671 (93h:60144)
[26] K. Majewski, Single class queueing networks with discrete and fluid customers on the time interval $\mathbb{R}$. *Queueing Systems* **36** (2000) 405–435.
[27] A.P. Markopoulou, F.A. Tobagi and M.J. Karam, Assessing the quality of voice communications over Internet backbones. *IEEE Transactions on Networking* **11** (2003) 747–760.
[28] A.A. Mogulskii, Large deviations for trajectories of multi-dimensional random walks. *Th. Prob. Appl.* **21** (1976) 300–315.
[29] S.V. Nagaev, Large deviations of sums of independent random variables. *Ann. Probab.* **7** (1979) 745–789.
[30] J. Neveu, Une generalisation des processus à accroissements positifs independants. *Abh. Math. Sem. Univ. Hamburg* **25** (1961) 36–61.

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press

[31] P. Ney and E. Nummelin, *Markov additive processes. I. Eigenvalue properties and limit theorems*. Ann. Probab. **15** (1987) 561–592.
[32] P. Ney and E. Nummelin, Markov additive processes II. Large deviations. *Ann. Probab.* **15** (1987) 593–609.
[33] S. Özekici and R. Soyer, Semi-Markov modulated Poisson process: probabilistic and statistical analysis. *Math. Methods Oper. Res.* **64** (2006) 125–144.
[34] A. Pacheco and N.U. Prabhu, *Markov-additive processes of arrivals*, Advances in queueing, Probab. Stochastics Ser., CRC, Boca Raton, FL (1995) 167–194.
[35] A. Puhalskii, Large deviation analysis of the single server queue. *Queueing Systems* **21** (1995) 5–66.
[36] A. Puhalskii and W. Whitt, Functional large deviation principles for first-passage-time proc esses. *Ann. Appl. Probab.* **7** (1997) 362–381.
[37] U. Rieder and N. Bäuerle, Portfolio optimization with unobservable Markov-modulated drift process. *J. Appl. Probab.* **42** (2005) 362–378.
[38] R.T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J. (1970).
[39] International Telecommunication Union, *Recommendation ITU-T P.59, Artificial Conversational Speech* (March 1993).

https://doi.org/10.1051/ps/2009010 Published online by Cambridge University Press