On large deviations in queuing systems *

Z.H. Li[†], E.A. Pechersky[‡]

[†] Department of Mathematics, Beijing Normal University,
Beijing 100875, China
[‡] Institute of Information Transmission Problems, Russian AS,
19, Bolshoi Karetnyi, GSP-4, Moscow 101447, Russia

Abstract

The main purpose of the article is to provide a simpler and more elementary alternative derivation of the large deviation principle for multi-dimensional compound Poisson processes defined on $[0, \infty)$. The result was originally obtained in [12], whose purpose was to establish a large deviation principle which may be used directly in the study of queuing systems and networks. Our new proof may be divided into two steps. In the first step, we obtain the large deviation principle for the processes relative to the vague topology from the finite dimensional Cramér's theorem by a projective limit argument. The result of this step is close to the one of Lynch and Sethuraman in [14], who considered one-dimensional processes defined on a finite interval. The second step is to extend the large deviation principle to the uniform-weak topology introduced in [12]. We do this by proving the exponential tightness of the processes under the uniform-weak topology and then applying the inverse contraction principle (see[2]).

*This work was supported in part by the National Natural Science Foundation of China (Grant 19361060) and the Russian Foundation for Basic Research (Grant 99-01-00003) and FAPESP (Grant 98/14795-3). Published in: *Resenhas do Instituto de Matematica e Estatistica da Universidade de Sao Paulo* **4** (1999), 163–182.

1 Introduction

The theory of large deviations has found wide applications in queuing systems and queuing networks (see [1, 2, 3, 4, 5, 6] and references there). According to the opinion of R. Dobrushin, there are mainly two research directions which can be considered as goals of the area. The first one is the investigation of the so-called "bottle neck" problems. Typically, this is to establish the fact that, under the condition of large delay of a message in the network, the message spent most of its time at a single node in the network. The second direction is to find explicit solutions of the large deviation problems for specific queuing systems. Those solutions can be used then instead of an exact analytical result. As it is well-known, there is no exact analytical results for more or less general networks. Even for the tandem queuing systems, the simplest one, there are only very cumbersome exact solutions in some particular cases. Therefore, more restricted but explicit results on the level of the large deviations would have practical interests.

Before the discussion, let us recall the definition of the large deviation principle (for example, see [2]). Let $\mathcal{X}$ be a topological space. Of course, $\mathcal{X}$ must possesses 'good' properties, for example, it is usually a polish space or a Banach space. However, we shall neglect this in the definition. Let $\{P_n, n = 1, 2, \ldots\}$ be a sequence of probability measures on the σ-algebra of the Borel subsets of this space that converges to a δ-measure concentrated at a point $x_0 \in \mathcal{X}$, that is

$$P_n \Rightarrow \delta_{x_0}.$$

Let

$$I : \mathcal{X} \to [0, \infty]$$

be a non-negative function possibly taking the infinite value.

One says that the *large deviations principle* holds for the sequence $\{P_n\}$ with the rate function I if

$$\liminf_{n\to\infty} \frac{1}{n} \ln P_n(\mathcal{B}) \geq - \inf_{\mathbf{x}\in\mathcal{B}^\circ} I(\mathbf{x})$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_n(\mathcal{B}) \leq - \inf_{\mathbf{x}\in\overline{\mathcal{B}}} I(\mathbf{x})$$

for each Borel set $\mathcal{B} \subseteq \mathcal{X}$. Here and in the sequel, $\overline{\mathcal{B}}$ denotes the closure of the set $\mathcal{B}$ and $\mathcal{B}^\circ$ denotes its interior.

The well-known Cramér's theorem described below is a typical example of the large deviations principle. Let $\mathcal{X} = \mathbb{R}$ and let P_n be the distribution of the average S_n/n, where $S_n = \sum_{i=1}^n \xi_i$ is the sum of the independent identically distributed random variables $\{\xi_i\}$. If $\mathsf{E}|\xi_1| < \infty$ than $P_n \Rightarrow \mathsf{E}\xi_1$. Assume the exponential decay of the distribution tail of ξ_i, namely, there exist constants $\theta_- < 0 < \theta_+$ such that

$$\varphi(\theta) = \mathsf{E}e^{\theta\xi_1} < \infty \tag{1}$$

for $\theta \in (\theta_-, \theta_+)$. The Cramér theorem states that the large deviation principle for the P_n is satisfied with the rate function

$$I(x) = \sup_\theta \{x\theta - \ln\varphi(\theta)\}.$$

We refer the reader to [2, 7, 8, 9] for details.

In queuing theory we need to consider large deviation principles for stochastic processes. Let us illustrate this by considering the classical queuing system having only one server and one input flow of messages to the server. Every message has a length. The server translates the messages with rate one, hence a message is treated by the server in a period of time equal to its length. The server works according to the first coming – first service (FCFS) discipline. It is assumed that there exists an infinite buffer where messages wait for their services if the server is occupied at their arrivals. The input flow can be described by a marked point process (η_n, ξ_n), where (η_n) is a random configuration of points on $\mathbb{R}$, which is interpreted as the times of message arrivals, and ξ_n is a mark assigned to the point η_n, which represents the length of the arrived message. In the following we assume that the input flow is a Poisson one, that is $\tau_n = \eta_{n+1} - \eta_n$ has the exponential distribution $\Pr(\tau_n > x) = \exp\{-\lambda x\}$ for a constant $\lambda > 0$; the random vectors (η_n, ξ_n) are independent and identically distributed; and the sequences (η_n) and (ξ_n) are independent. We assume further that $\lambda\varphi'(0) < 1$, so the system has a steady state. Let us consider a system in its steady state. The processes of interests are the length $\nu(t)$ of a queue and the time delay $\omega(t)$ of a message in the system if it arrives to the system at the moment t. The processes $\nu(t)$ and $\omega(t)$ are determined by the input flow in a unique way. The large deviation principle for those process is not an easy problem. Observe that both $\nu(t)$ and $\omega(t)$ are Markov processes if the distribution of ξ_1 is exponential (see, however, [10] and [11]).

It is natural to think that a designer of the queuing systems needs to estimate the probabilities

$$\Pr(\nu(t) > b) \quad \text{and} \quad \Pr(\omega(t) > a)$$

at a *fixed* moment t. Because the system is studied in the steady state we can take $t = 0$. It is possible to express the functionals $\nu(0)$ and $\omega(0)$ in terms of the process (η_n, ξ_n). For example,

$$\omega(0) = \sup_{t \geq 0} \{\zeta(t) - t\},$$

where $\zeta(t)$ is a compound Poisson process in terms of (η_n, ξ_n) as

$$\zeta(t) = \sum_{i:\, 0 \leq \eta_i < t} \xi_i.$$

Therefore, it is enough to have the large deviation principle for the process $\zeta(t)$. This program was realized in [12].

Before the work of [12], many works have been devoted to the principles of large deviations for processes with independent increments and other similar classes of processes (for example, see [13], [14], [15], [16]). However, none of them can be applied directly to the above problems. The first difficulty is connected with the usually imposed hypothesis that the exponential moment of jumps $\mathsf{E}e^{\theta\xi}$ of the compound Poisson process are finite for all θ. This excludes the important special case where ξ has an exponential distribution, which corresponds to the most important exponential service time queuing system. Including the case where $\mathsf{E}e^{\theta\xi}$ is finite for only small enough θ essentially complicates the study of the large deviations. In the case when all exponential moments exist a large increment of the process can arise only as a result of a cumulative contribution of many small jumps, while in the case of exponential distribution such a increment can also arise as a result of one big jump. It turns out that the formula for the rate function, which is well-known for the case of the processes with all finite exponential moments, requires an essential modification in the case of infinite moments.

The interesting large deviation principle for compound Poisson processes of Lynch and Sethuraman [14] includes the case of infinite exponential moments. But they only considered scalar-valued processes defined on a finite interval. In the study of queuing systems, we need to consider vector-valued processes defined on the half-line. (Observe that $\omega(0)$ is a functional of the process on the half-line.) Although the large deviation principles on finite intervals may be extended to their half-line versions using the projective limit arguments, the resulted projective limit topology on the trajectory space are too weak for the applications. For example, if we consider the event $(\omega(0) > x)$ for a positive x, then the set of trajectories corresponding to this event is

$$\mathcal{A} = \left\{ \mathbf{x}(t) : \sup_{t \geq 0} \{\mathbf{x}(t) - t\} > x \right\}.$$

However, it turns out that the closure $\overline{\mathcal{A}}$ of the set $\mathcal{A}$ in the projective limit topology includes each of the trajectories

$$\mathbf{x}(t) = at, \quad t \in [0, \infty),$$

where $a \in \mathbb{R}^1$. Indeed, the sequence

$$\mathbf{x}_n(t) = \begin{cases} at, & \text{if } t \leq n, \\ 2(t-n) + an, & \text{if } T > n \end{cases}$$

from $\mathcal{A}$ converges to $\mathbf{x}(t) = at$ in the projective limit. Therefore, applications of large deviation principles for the projective limit topology only give trivial estimates.

A new topology, the uniform-weak topology, on the trajectory space was introduced in [12], which overcomes the shortcomings mentioned above. (We shall review the definition of this topology in the next section.) In [12] a large deviation principle was established for vector-valued compound Poisson processes on $[0, \infty)$ relative to the uniform-weak topology. An application of this large deviation principle to the simplest network, the tandem system, was given in [5]. It was shown there that the bottle neck effect holds in the tandem system on the level of large fluctuations of the delay. In [6] the large deviations for the two dimensional functional $(\nu(0), \omega(0))$ was obtained. The large deviation principle for a two dimensional compound Poisson process with dependent components was involved in this investigation.

Since the proof of the large deviation principle in [12] was sophisticated and based on a general large deviation principle on abstract vector spaces, we think it is of interest to provide a simpler and more elementary alternative derivation of the result from the viewpoint of applications. This is the main purpose of the present paper. Since only compound Poisson processes *with non-decreasing paths* are used in the study of queuing systems, we shall restrict to this particular type of processes, which simplifies considerably the proof. Our proof may be divided into two steps. In the first step, we obtain the large deviation principle for the compound Poisson process relative to the vague topology from the finite dimensional Cramér's theorem by a projective limit argument. In this step we calculate the rate function. The difference from [14] is that Lynch and Sethuraman considered one-dimensional processes on a finite interval while we consider multi-dimensional processes on the half line. The second step is to extend the large deviation principle from the vague topology to the uniform-weak one using the inverse contraction principle. This step enlarges the class of Borel sets keeping the same the rate function.

2 Large deviation principle for compound Poisson processes on $[0, \infty)$

Now we specialize the three objects introduced in the definition of a large deviation principle: the topological space $\mathcal{X}$, the sequence of measures $\{P_n\}$ and the rate function I.

We start with the space $\mathcal{X}$. Let $\mathcal{X}$ be the space of functions $\mathbf{x} : (-\infty, \infty) \to \mathbb{R}^r$ with the following three properties:

1) For any $\mathbf{x} \in \mathcal{X}$ and any $t < 0$ we have $\mathbf{x}(t) = 0$.

2) All the functions $\mathbf{x} \in \mathcal{X}$ are non-decreasingly monotone.

3) The functions $\mathbf{x} \in \mathcal{X}$ are right-continuous at each point $t \in \mathbb{R}$, i.e.,

$$\mathbf{x}(t) = \mathbf{x}(t + 0) = \lim_{u \downarrow t} \mathbf{x}(u).$$

4) The limits

$$v(\mathbf{x}) = \lim_{t \to \infty} \frac{\mathbf{x}(t)}{1 + t}$$

exist and are finite.

The condition 1) is rather formal. It is convenient way to include a jump of $\mathbf{x}$ at 0.

There is a one-to-one correspondence between the functions $\mathbf{x} \in \mathcal{X}$ and positive $\mathbb{R}^r$-valued measures $\mu_{\mathbf{x}}$ on $[0, \infty)$. This correspondence is defined by the relation

$$\mathbf{x}(t) = \mu_{\mathbf{x}}([0, t]), \quad t \geq 0.$$

Let Φ be the set of all continuous $\mathbb{R}^r$-valued functions $\phi(t)$, $t \in \mathbb{R}$, with compact support, i.e. it vanishes out of a finite interval $[-T_\phi, T_\phi]$. For $\phi \in \Phi$ and $\mathbf{x} \in \mathcal{X}$ we let

$$J_\phi(\mathbf{x}) = \langle \phi, \mathbf{x} \rangle = \int_0^\infty \phi(t)\mathbf{x}(t)\, dt.$$

Here and in the following ab is the inner product of vectors $a, b \in \mathbb{R}^r$. It is clear that, for any fixed function $\phi \in \Phi$, this defines a linear functional J_ϕ on the space $\mathcal{X}$. To each function $\phi \in \Phi$ we associate a function

$$\hat{\phi}(t) = \int_t^\infty \phi(u)\, du. \tag{2}$$

The function $\hat{\phi}$ is continuously differentiable and vanishes for large enough $t \geq 0$. Then

$$J_\phi(\mathbf{x}) = \int_0^\infty \hat{\phi}(t)\mu_{\mathbf{x}}(dt). \tag{3}$$

For any number $T \geq 0$ we define the *shift operator* S_T by

$$S_T\phi(t) = \phi(t - T), \quad t \in \mathbb{R}^1,\ \phi \in \Phi.$$

The topology on $\mathcal{X}$ is defined by the system of pseudometrics

$$\rho_\phi(\mathbf{x}, \mathbf{y}) = \sup_{t \geq 0} \left\{ \frac{1}{1+t} |J_{S_n\phi}(\mathbf{x} - \mathbf{y})| \right\}, \quad \phi \in \Phi.$$

That is, a sequence $\{\mathbf{x}_N \in \mathcal{X},\ N = 1, 2, \ldots\}$ converges to $\mathbf{x} \in \mathcal{X}$ if and only if

$$\lim_{N \to \infty} \rho_\phi(\mathbf{x}, \mathbf{x}_N) = 0 \quad \text{for all} \quad \phi \in \Phi. \tag{4}$$

We shall call this topology the *uniformly-weak topology*.

Next we describe the probability measures $\{P_n\}$. Let us recall a description of a compound Poisson process $\zeta(t)$. Suppose that π be a positive measure on the space $\mathbb{R}^r$ such that

$$\int_{\mathbb{R}^r} |y|\pi(dy) < \infty. \tag{5}$$

A probability measure P^π on Borel subsets of the space $\mathcal{X}$ is the distribution of a compound Poisson process $\zeta(t)$ with jump measure π if for any function

$\phi \in \Phi$ the characteristic function is

$$\mathsf{E}e^{i\langle\phi,\zeta\rangle} = \int_{\mathcal{X}} \exp\{iJ_\phi(x)\}\, P_\pi(dx) = \exp\left\{\int_0^\infty \int_{\mathbb{R}^r} (\exp\{iy\hat{\phi}(t)\} - 1)\pi(dy)dt\right\} \tag{6}$$

(see the notation (2)). Heuristically this means that we are considering a time-homogeneous Poisson process such that the probability for a jump with size $y \in A$ to happen in the time interval dt is equal to $\pi(A)dt$ if $\pi(A) < \infty$. It follows from the definition (6) that

$$v(\mathbf{x}) = m,$$

with the P^π-probability 1, where

$$m = \int_{\mathbb{R}^r} y\pi(dy)$$

is the mean value of π. For the considered case of non-decreasing paths the measure π is concentrated in $\mathbb{R}^r_+$.

Let $\mathbf{F}_n : \mathcal{X} \to \mathcal{X}$ be the transformation

$$\mathbf{x}(t) \to \mathbf{x}_n(t) = \frac{1}{n}\mathbf{x}(nt). \tag{7}$$

It is easy to check that the conditions 1) – 4) included in the definition of the space $\mathcal{X}$ are valid for the function $\mathbf{x}_n(t)$ if they are valid for $\mathbf{x}(t)$. Let P^π_n be the measure on $\mathcal{X}$ induced by the transformation $\mathbf{F}_n$ from the measure P^π. It is easy to understand that P^π_n defines again a compound Poisson process $\zeta_n(t)$ with the jump measure

$$\pi_n(A) = n\pi(nA).$$

Obviously we have

$$\zeta_n(t) = \frac{\zeta(nt)}{n}. \tag{8}$$

Now let us define the rate function I. Suppose that for some $a > 0$ we have

$$\int_{\mathbb{R}^r} (\mathrm{e}^{a|y|} - 1)\pi(dy) < \infty. \tag{9}$$

This inequality implies the condition (5). Let

$$q(\theta) = \int_{\mathbb{R}^r} (\mathrm{e}^{\theta y} - 1)\pi(dy), \quad \theta \in \mathbb{R}^r, \tag{10}$$

and let Θ_π be the set of points $\theta \in \mathbb{R}^r$ for which $q(\theta) < \infty$. It is easy to check that $q(\theta)$ is a convex function of $\theta \in \mathbb{R}^r$ and so Θ_π is a convex set.

Let

$$\Lambda_{\rm a}(x) = \sup_{\theta\in\Theta_\pi} \{\theta x - q(\theta)\}, \quad x \in \mathbb{R}^r. \tag{11}$$

The function $\Lambda_{\rm a}(x)$ is called the *Legendre transformation* of the function q. It is a convex function of x with values in $[0, \infty]$. (Observe that $\theta x - q(\theta) = 0$ if $\theta = 0$.) Let Θ_π° be a set of all inner points of the set Θ_π, which is non-empty because of the condition (9). It is clear that $q(\theta)$ is smooth in the domain Θ_π°. If for some $x \in \mathbb{R}^r$ there exists $\theta_x \in \Theta_\pi^\circ$ such that the value of the gradient

$$\nabla q(\theta_x) = x,$$

then

$$\Lambda_{\rm a}(x) = \theta_x x - q(\theta_x)$$

(see [17], §26). It is clear that $\nabla q(0) = m$ and so

$$\Lambda_{\rm a}(m) = 0. \tag{12}$$

For an absolutely continuous function $\mathbf{x}_{\rm a} \in \mathcal{X}$, i.e.,

$$\mathbf{x}_{\rm a}(t) = \int_0^\infty \dot{\mathbf{x}}_{\rm a}(u)du,$$

we set

$$I_{\rm a}(\mathbf{x}_{\rm a}) = \int_0^\infty \Lambda_{\rm a}(\dot{\mathbf{x}}_{\rm a}(u))du. \tag{13}$$

(The integral is meaningful because $\Lambda_{\rm a} \geq 0$.)

Let

$$\Lambda_s(x) = \sup_{\theta\in\Theta_\pi} \theta x. \tag{14}$$

Then we have

$$\Lambda_s(x) = \lim_{\gamma\to\infty} \frac{1}{\gamma}\Lambda_{\rm a}(\gamma x), \quad x \in \mathbb{R}^r.$$

In order to see this, let $U \subset \mathbb{R}^r$ be the ball centered at 0 with radius 1. If $\theta \notin \Theta_\pi$ then $\theta x - q(\theta) = -\infty$. Therefore

$$\limsup_{\gamma\to\infty} \frac{1}{\gamma}\Lambda_{\rm a}(\gamma x)$$
$$\leq \limsup_{\gamma\to\infty} \left[\sup_{\theta\in\Theta_\pi} \left\{\theta x - \frac{1}{\gamma}\theta \int_U y\pi(dy)\right\} - \frac{1}{\gamma}\int_{U^c} \pi(dy)\right] = \Lambda_s(x).$$

On the other hand, for any fixed $\theta \in \Theta_\pi$ and x we have

$$\gamma^{-1}(\theta\gamma x - q(\theta)) \to \theta x$$

as $\gamma \to \infty$. Therefore

$$\liminf_{\gamma\to\infty} \frac{1}{\gamma}\Lambda_a(\gamma x) \geq \theta x, \tag{15}$$

as desired. We note also that $\Lambda_s(x)$ is a convex non-negative function of $x \in \mathbb{R}^r$ (see [17], §13). It is linear on the ray $\{\lambda x, 0 \leq \lambda < \infty\}$ for any $x \in \mathbb{R}^r$. If $r = 1$, then

$$\Lambda_s(x) = \begin{cases} x \sup\{\theta : \theta \in \Theta_\pi\}, & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x = 0, \\ \infty, & \text{if} \quad x < 0. \end{cases} \tag{16}$$

We introduce the system $\mathcal{P}$ of all finite partitions $\Pi = \{-\infty < t_0 < t_1 < \cdots < t_n < \infty\}$, $n = 1, 2, \ldots$. Let $\mathbf{x}_s$ is singular, i.e. a function such that the corresponding measure $\mu_{\mathbf{x}_s}$ is singular with respect to the Lebesgue measure. We define

$$I_s^\Pi(\mathbf{x}_s) = \sum_{k=1}^{n} \Lambda_s(\mathbf{x}_s(t_k) - \mathbf{x}_s(t_{k-1}))$$

then

$$I_s(\mathbf{x}_s) = \sup_{\Pi\in\mathcal{P}} \left\{ I_s^\Pi(\mathbf{x}_s) \right\}. \tag{17}$$

Any function $\mathbf{x} \in \mathcal{X}$ can be represented in a unique way as

$$\mathbf{x} = \mathbf{x}_a + \mathbf{x}_s,$$

where $\mathbf{x}_a$ is an absolutely continuous function, and x_s is singular. The rate function I is the sum

$$I(\mathbf{x}) = I_a(\mathbf{x}_a) + I_s(\mathbf{x}_s). \tag{18}$$

We shall say that a partition $\Pi' = \{-\infty < t'_0 < t'_1 < \ldots < t'_{n'} < \infty\}$ is a *subpartition* of the partition Π if each point t'_i coincides with one of the points t_k. It follows from non-negativity and subadditivity of the function Λ_s that, if Π' is a subpartition of Π, then

$$I_s^{\Pi'}(\mathbf{x}_s) \leq I_s^\Pi(\mathbf{x}_s).$$

Hence we can also interpret $I_s(\mathbf{x}_s)$ as the limit of $I_s^\Pi(\mathbf{x}_s)$ with respect to the partial ordering of the set $\mathcal{P}$ defined by the subpartitions.

The main theorem of this paper is the following

Theorem 1 ([12]) *If condition (9) is true then the sequence of the probability measures* $\{P_n^\pi, n = 1, 2, \ldots\}$ *satisfies the large deviations principle with the rate function* I *defined in (18), (13), and (17).*

3 Proof of the theorem

As mentioned before, our proof may be divided into two steps. In the first step, we obtain the large deviation principle for the process relative to the vague topology from the finite dimensional Cramér's theorem by a projective limit argument. The second step is to extend the large deviation principle from the vague topology to the uniform-weak one using the inverse contraction principle.

3.1 Large deviation principle for the compound Poisson process under vague topology

Recall that P_n^π denotes the distribution on $\mathcal{X}$ of the processes ζ_n (see (8)). Given a partition $\Pi = \{-\infty < t_0 < t_1 < \cdots < t_m < \infty\}$ let

$$I^\Pi(\mathbf{x}) = \sum_{i=1}^{m} (t_i - t_{i-1}) \Lambda_a([\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})]/(t_i - t_{i-1})). \tag{19}$$

In the next proposition we define the rate function on $\mathcal{X}$ by

$$I(\mathbf{x}) = \sup_{\Pi \in \mathcal{P}} I^\Pi(\mathbf{x}). \tag{20}$$

The one-dimensional version of the following large deviation principle was proved by Lynch and Sethuraman [14].

Proposition 2 *The sequence* $\{P_n^\pi\}$ *satisfies the large deviation principle with rate function defined by (20) and (19).*

Proof. The random vector $\zeta_n(t) - \zeta_n(s)$ has the same distribution as $\frac{1}{n}\sum_{k=1}^{n} \zeta^k(t-s)$, where $\zeta^k(u)$ $k = 1, ..., n$, is a sequence of independent identically distributed processes having the distribution coinciding with $\zeta(u)$.

By Cramér's theorem, the sequence $\zeta_n(t) - \zeta_n(s)$ satisfies the large deviation principle with regular rate function

$$(t-s)\Lambda_a(x/(t-s)) = \sup_{\theta \in \Theta_\pi} \{\theta x - (t-s)q(\theta)\}, \quad x \in \mathbb{R}^r.$$

Let $\Pi = \{-\infty < t_0 < ... < t_m < \infty\}$ be a partition. By the independent increments property, the distributions of $(\mathbf{x}(t_1)-\mathbf{x}(t_0), \ldots, \mathbf{x}(t_m)-\mathbf{x}(t_{m-1}))$ under P_n^π, $n = 1, 2, \ldots$, satisfies the large deviation principle with regular rate function

$$J^\Pi(x_1, \ldots, x_m) := \sum_{i=1}^{m} (t_i - t_{i-1}) \Lambda_{\mathrm{a}}(x_i/(t_i - t_{i-1})).$$

Note that the map $\mathbb{R}^{rm} \to \mathbb{R}^{rm}$ defined by $(\mathbf{x}(t_1)-\mathbf{x}(t_1), ..., \mathbf{x}(t_m)-\mathbf{x}(t_{m-1})) \mapsto (\mathbf{x}(t_1), ..., \mathbf{x}(t_m))$ is continuous. By the contraction principle (see [2, Theorem 4.2.1]), the distributions of $(\mathbf{x}(t_1), \ldots, \mathbf{x}(t_m))$ under P_n^π, $n = 1, 2, \ldots$, satisfies the large deviation principle with regular rate function I^Π defined by (19). Since the pointwise convergence in $\mathcal{X}$ implies the vague convergence, the desired large deviation principle follows by a projective limit argument (see e.g. [18] or [2, Theorem 4.7.1]). □

Proposition 3 *Let I be defined by (20) and (19). Then for any $\mathbf{x} \in \mathcal{X}$ we have*

$$I(\mathbf{x}) = \sup_{f \in B[0,\infty)} \left\{ \int_0^\infty f(t) \mu_{\mathbf{x}}(dt) - \int_0^\infty q(f(t)) dt \right\}, \tag{21}$$

where $B[0, \infty)$ is the set of bounded Borel functions on $[0, \infty)$ taking its values in Θ_π. The equality remains true when $B[0, \infty)$ is replaced by $C[0, \infty)$ or $D[0, \infty)$, where $C[0, \infty) = \{$ continuous functions in $B[0, \infty)\}$ and $D[0, \infty)$ $= \{$ piecewise constant functions in $B[0, \infty)\}$.

Proof. For any $\Pi \in \mathcal{P}$ let $D_0^\Pi[0, \infty)$ be the subset of $D[0, \infty)$ consisting of functions which have bounded supports and are constant on each partition interval of Π. From (19) it is not hard to see that

$$I^\Pi(\mathbf{x}) = \sup_{f \in D_0^\Pi[0,\infty)} \left\{ \int_0^\infty [f(t)\dot{\mathbf{x}}^\Pi(t) - q(f(t))] dt \right\}, \tag{22}$$

where

$$\mathbf{x}^\Pi(t) = \begin{cases} \mathbf{x}(t_{i-1}) + \frac{\mathbf{x}(t_i)-\mathbf{x}(t_{i-1})}{t_i - t_{i-1}}(t - t_{i-1}), & \text{if } t_{i-1} \le t \le t_i, \\ \mathbf{x}(t_m) + mt, & \text{if } t > t_m. \end{cases}$$

It follows from (20) that

$$I(\mathbf{x}) = \sup_{f \in D[0,\infty)} \sup_{\Pi \in \mathcal{P}} \left\{ \int_0^\infty [f(t)\dot{\mathbf{x}}^\Pi(t) - q(f(t))] dt \right\}. \tag{23}$$

For any $f \in D[0,\infty)$ with bounded support we have $f \in D_0^\Pi[0,\infty)$ for some $\Pi \in \mathcal{P}$, so

$$\int_0^\infty f(t)\mu_{\mathbf{x}}(dt) = \int_0^\infty f(t)\dot{\mathbf{x}}^\Pi(t)dt. \tag{24}$$

In view of (23) and (24) we have

$$I(\mathbf{x}) \geq \sup_{f\in D[0,\infty)} \left\{ \int_0^\infty f(t)\mu_{\mathbf{x}}(dt) - \int_0^\infty q(f(t))dt \right\}. \tag{25}$$

Using a monotone class argument one sees that $D[0,\infty)$ is dense in $B[0,\infty)$ by pointwise convergence. Therefore, (25) yields

$$I(\mathbf{x}) \geq \sup_{f\in B[0,\infty)} \left\{ \int_0^\infty f(t)\mu_{\mathbf{x}}(dt) - \int_0^\infty q(f(t))dt \right\}. \tag{26}$$

On the other hand, by (20) for any $\eta < I(\mathbf{x})$ there is $\Pi \in \mathcal{P}$ satisfying $\eta < I^\Pi(\mathbf{x})$. By (22), we can find $f \in D_0^\Pi[0,\infty)$ such that

$$\eta < \int_0^\infty [f(t)\dot{\mathbf{x}}^\Pi(t) - q(f(t))]dt = \int_0^\infty f(t)\mu_{\mathbf{x}^\Pi}(dt) - \int_0^\infty q(f(t))dt.$$

It follows that

$$I(\mathbf{x}) \leq \sup_{f\in D[0,\infty)} \left\{ \int_0^\infty f(t)\mu_{\mathbf{x}}(dt) - \int_0^\infty q(f(t))dt \right\}.$$

Since each $f \in D[0,\infty)$ can be approximated by a sequence $\{f_n\} \subseteq C[0,\infty)$, the inequality also holds when $D[0,\infty)$ is replaced by $C[0,\infty)$. These and (26) yield the desired equalities. □

Proposition 4 *For any* $\mathbf{x} = \mathbf{x}_a + \mathbf{x}_s \in \mathcal{X}$ *we have*

$$I(\mathbf{x}) = I_a(\mathbf{x}_a) + \widetilde{I}_s(\mathbf{x}_s). \tag{27}$$

where

$$\widetilde{I}_s(\mathbf{x}_s) = \sup_{f\in B[0,\infty)} \left\{ \int_0^\infty f(t)\mu_{\mathbf{x}_s}(dt) \right\}. \tag{28}$$

Proof. By Proposition 3 we at least have

$$I(\mathbf{x}) \le I_{\mathrm{a}}(\mathbf{x}_{\mathrm{a}}) + \widetilde{I}_s(\mathbf{x}_s). \tag{29}$$

Moreover, there is a sequence $\{f_n\} \subseteq B[0,\infty)$ such that

$$\int_0^\infty f_n(t)\mu_{\mathbf{x}_{\mathrm{a}}}(dt) - \int_0^\infty q(f_n(t))dt \to I_{\mathrm{a}}(\mathbf{x}_{\mathrm{a}})$$

as $n \to \infty$. Similarly, by (28) there is $\{g_n\} \subseteq B[0,\infty)$ such that

$$\int_0^\infty g_n(t)\mu_{\mathbf{x}_s}(dt) \to \widetilde{I}_s(\mathbf{x}_s)$$

as $n \to \infty$. Let $F \subset [0,\infty)$ be of zero Lebesgue measure and $\mu_{\mathbf{x}_s}(F) = \mu_{\mathbf{x}_s}[0,\infty)$. Define the sequence $\{h_n\} \subseteq B$ by

$$h_n(t) = f_n(t)1_{F^c}(t) + g_n(t)1_F(t).$$

It is easy to see that

$$\int_0^\infty h_n(t)\mu_{\mathbf{x}}(dt) - \int_0^\infty q(h_n(t))dt \to I(\mathbf{x}_{\mathrm{a}}) + \widetilde{I}_s(\mathbf{x}_s).$$

Then (27) follows from (21) and (29). ☐

It is simply to check that

$$I_s(\mathbf{x}_s) = \widetilde{I}_s(\mathbf{x}_s).$$

Now we have obtained the desired large deviation principle under the vague topology.

3.2 Inverse contraction principle

Suppose that $\{P_n\}$ is a sequence of the probability measures on a topological space $\mathcal{Y}$ such that $P_n \Rightarrow \delta_x$ for $x \in \mathcal{Y}$. We say $\{P_n\}$ is *exponentially tight* if for any $\varepsilon > 0$ there is a compact set $\mathcal{K} \subset \mathcal{Y}$ such that $P_n(K^c) \le \varepsilon^n$.

To extend the large deviation principle to the uniform-weak topology we appeal to the following

Theorem 5 ([2, Corollary 4.2.10]) *Let a set $\mathcal{X}$ be equipped with two Hausdorff topologies τ_1 and τ_2, where τ_2 is coarser than τ_1. Assume that a sequence of probability measures $\{P_n\}$ satisfies the large deviation principle with rate function $I : \mathcal{X} \to [0, \infty]$ under the topology τ_2. If $\{P_n\}$ is exponentially tight with respect to the topology τ_1, then the large deviation principle holds for $\{P_n\}$ with the same rate function I relative to τ_1.*

Since the uniform-weak topology is finer than the vague topology, in order to finish the proof of Theorem 1 we need only to show that $\{P_n^\pi\}$ (see the section 2) is exponentially tight in the uniform-weak topology. In the following two subsections, we shall give a description of a class of compact subsets of $\mathcal{X}$ and use it to prove the exponential tightness.

3.2.1 Compact sets

In this subsections, we describe a class of compact subsets of $\mathcal{X}$.

Lemma 6 *If $\mathbf{x}_n \to \mathbf{x}$ in $\mathcal{X}$ in the uniform weak topology, then we have*

$$\lim_{n\to\infty} \langle \mathbf{x}_n, \phi \rangle = \langle \mathbf{x}, \phi \rangle, \quad \phi \in \Phi, \tag{30}$$

and

$$\lim_{n\to\infty} v(\mathbf{x}_n) = v(\mathbf{x}). \tag{31}$$

Under the additional condition

$$\lim_{t\to\infty} \sup_n \left| \frac{\mathbf{x}_n(t)}{1+t} - v(\mathbf{x}_n) \right| = 0, \tag{32}$$

in order that $\mathbf{x}_n \to \mathbf{x}$ in the uniform weak topology it is necessary and sufficient that (30) and (31) hold.

Proof. First observe that for any $\phi \in \Phi$ and $\mathbf{x} \in \mathcal{X}$ we have

$$\lim_{k\to\infty} \left| \frac{\langle \mathbf{x}, S_k\phi \rangle}{1+k} - v(\mathbf{x}) \int_{-\infty}^{\infty} \phi(t)dt \right| = 0. \tag{33}$$

Suppose that $\mathbf{x}_n \to \mathbf{x}$ in $\mathcal{X}$ in the uniform weak topology. The relation (30) follows from (4) in an evident way. By (33), we have

$$\lim_{k\to\infty} \frac{1}{1+k} \langle \mathbf{x}_n - \mathbf{x}, S_k\phi \rangle = (v(\mathbf{x}_n) - v(\mathbf{x})) \int_{-\infty}^{\infty} \phi(t)dt.$$

From this and the definition of ρ_ϕ it follows that

$$\rho_\phi(\mathbf{x}_n, \mathbf{x}) \geq \left| (v(\mathbf{x}_n) - v(\mathbf{x})) \int_{-\infty}^{\infty} \phi(t)\, dt \right|.$$

Since $\lim_{n\to\infty} \rho_\phi(\mathbf{x}_n, \mathbf{x}) = 0$ for all $\phi \in \Phi$, we get (31).

Now we assume that the conditions (30), (31) and (32) are satisfied. Observe that

$$\left| \frac{\langle \mathbf{x}_n, S_k\phi \rangle}{1+k} - v(\mathbf{x}_n) \int_{-\infty}^{\infty} \phi(t) dt \right| = \left| \int_{-\infty}^{\infty} \phi(t) \left(\frac{\mathbf{x}_n(t+k)}{1+k} - v(\mathbf{x}_n) \right) dt \right|$$
$$\leq= \int_{-\infty}^{\infty} |\phi(t)| \left| \frac{\mathbf{x}_n(t+k)}{1+t+k} - v(\mathbf{x}_n) \right| dt + \int_{-\infty}^{\infty} |\phi(t)| \frac{|\mathbf{x}_n(t+k)t|}{(1+k)|1+t+k|} dt. \quad (34)$$

Both terms on the right hand side of (34) goes to zero uniformly in $n = 1, 2, \dots$ as $k \to \infty$. Then we have proved that

$$\lim_{k\to\infty} \sup_n \left| \frac{\langle \mathbf{x}_n, S_k\phi \rangle}{1+k} - v(\mathbf{x}_n) \int_{-\infty}^{\infty} \phi(t) dt \right| = 0. \quad (35)$$

It follows from (33) and (35) that for any $\varepsilon > 0$ there exists a value $N = N(\varepsilon, \phi)$ such that

$$\left| \frac{\langle \mathbf{x}_n - \mathbf{x}, S_k\phi \rangle}{1+k} \right| \leq \varepsilon + |v(\mathbf{x}_n) - v(\mathbf{x})| \int_{-\infty}^{\infty} |\phi(t)| dt$$

for all $n \geq 1$ and $k \geq N$. The condition (31) enables us to state

$$\sup_{k \geq N} \left| \frac{\langle \mathbf{x}_n - \mathbf{x}, S_k\phi \rangle}{1+k} \right| \leq 2\varepsilon$$

for large enough $n \geq 0$. Then we see that the conditions (30), (31) and (32) imply the convergence $\mathbf{x}_n \to \mathbf{x}$ in the uniform weak topology. □

Lemma 7 *Let C and D be right continuous, nonnegative functions on $[0, \infty)$. Suppose that C is increasing, D is decreasing and $D(t) \to 0$ as $t \to \infty$. Let $\mathcal{K} = \mathcal{K}(C, D)$ be the set of functions $\mathbf{x} \in \mathcal{X}$ such that $\mathbf{x}(t) \leq C(t)$ and*

$$\left| \frac{\mathbf{x}(t)}{1+t} - v(\mathbf{x}) \right| \leq D(t) \quad (36)$$

for all $t \in [0, \infty)$. Then $\mathcal{K}$ is a metrizable, compact subset of $\mathcal{X}$.

Proof. Let d be a metric on $\mathcal{X}$ that agrees with the vague topology and let

$$\rho(\mathbf{x},\mathbf{y}) = d(\mathbf{x},\mathbf{y}) + |v(\mathbf{x}) - v(\mathbf{y})|, \quad \mathbf{x},\mathbf{y} \in \mathcal{K}.$$

By Lemma 6 one may see that ρ is a metric on $\mathcal{K}$ that agrees with the uniform weak topology, that is, $\mathcal{K}$ is a metrizable space. Consequently, we need only to show that $\mathcal{K}$ is closed and sequentially compact. Consider a sequence $\{\mathbf{x}_n\} \subseteq \mathcal{K}$ such that $\mathbf{x}_n \to \mathbf{x}$ in the uniform weak topology in $\mathcal{X}$. By Lemma 6 we have $v(\mathbf{x}_n) \to v(\mathbf{x})$ and $\mathbf{x}_n \to \mathbf{x}$ vaguely. Then $\mathbf{x}_n(t) \to \mathbf{x}(t)$ for all continuity points $t \geq 0$ of $\mathbf{x}$. It follows that $\mathbf{x}(t) \leq C(t)$, first for continuity points of $\mathbf{x}$ and then for all $t \geq 0$ by the right continuity. Similarly we get (36). Therefore, $\mathcal{K}$ is a closed subset of $\mathcal{X}$.

Now let $\{\mathbf{x}_n\} \subseteq \mathcal{K}$. Then the sequence $\{\mathbf{x}_n(t)\}$ is bounded for each $t \geq 0$. By the definition of $\mathcal{K}$ we have $|v(\mathbf{x}_n)| \leq C(0) + D(0)$. It follows that the sequence $\{v(\mathbf{x}_n)\}$ is also bounded. Using a diagonal procedure we can find a subsequence $\{n_k\}$ such that $v(\mathbf{x}_{n_k}) \to$ some $v \in [0,\infty)$ and $\mathbf{x}_{n_k} \to$ some $\mathbf{x} \in \mathcal{X}$ vaguely as $k \to \infty$. Then (36) holds clearly. Therefore, $\mathbf{x} \in \mathcal{X}$ and $\mathcal{K}$ is sequentially compact. ☐

3.2.2 Exponential tightness

We prove in this subsection the exponential tightness for the processes ζ_n (see (8)) having the probability distribution $\{P_n^\pi\}$, from which the main theorem follows.

Proposition 8 *For any $\epsilon > 0$ there exists a compact $\mathcal{K}(\epsilon) \subset \mathcal{X}$ such that for any n the probability*

$$P_n^\pi\left(\mathbf{x} \notin \mathcal{K}(\epsilon)\right) = \Pr\left(\zeta_n \notin \mathcal{K}(\epsilon)\right) \leq \epsilon^n.$$

Proof follows from the following lemmas. ☐

Lemma 9 *For any $\varepsilon > 0$ there exists a right continuous, nonnegative, increasing function C_ε on $[0,\infty)$ such that*

$$\Pr\left\{\zeta_n(t) > C_\varepsilon(t) \text{ for some } t \geq 0\right\} \leq \varepsilon^n/2 \tag{37}$$

for all $n = 1, 2, \ldots$.

Proof. Let $\eta > 0$ be the constant such that $\eta \le a$ (see (9)). By Chebyshev's inequality, for $l > 0$ we have

$$\Pr\{\zeta(nt) > nl\} \le \exp\{ntq(\eta, \dots, \eta) - n\eta l\}. \tag{38}$$

For any $K \ge 0$ there is a value $\alpha = \alpha(K)$ so large that $q(\eta, \dots, \eta) - \alpha\eta \le -K$. Letting $l = \alpha t$ in (38) we see that

$$\Pr\{\zeta_n(t) > \alpha t\} \le \exp\{-nKt\}.$$

Consequently,

$$\sum_{k=1}^{\infty} \Pr\{\zeta_n(k) > \alpha k\} \le \frac{\exp\{-nK\}}{1 - \exp\{-nK\}}. \tag{39}$$

But the path of $\zeta_n(t) : t \ge 0$ is non-decreasing, so $\zeta_n(t) > \alpha(t+1)$ for some $t \ge 0$ implies $\zeta_n(k) > \alpha k$ for some integer $k \ge 1$. By (39) it follows that

$$\Pr\left\{\zeta_n(t) > \alpha(K)(t+1) \text{ for some } t \ge 0\right\} \le \frac{\exp\{-nK\}}{1 - \exp\{-nK\}}. \tag{40}$$

For any $\varepsilon > 0$ we choose the value $K = K(\varepsilon) \ge 0$ such that $e^{-K} \le \min\{1/2, \varepsilon/4\}$ and let $C_\varepsilon(t) = \alpha(K)(t+1)$. Then we get (37) from (40). $\square$

Lemma 10 *For any integers $n, p \ge 1$ we have*

$$\sum_{k=p}^{\infty} e^{-n\sqrt{k}} \le \frac{2}{n}(\sqrt{p-1} + 1)e^{-n\sqrt{p-1}}. \tag{41}$$

Proof The inequality follows as we observe

$$\sum_{k=p}^{\infty} e^{-n\sqrt{k}} \le \int_{p-1}^{\infty} e^{-n\sqrt{x}} dx = 2\int_{\sqrt{p-1}}^{\infty} ye^{-ny} dy$$

and compute the value on the right side. $\square$

Lemma 11 *For any $\varepsilon > 0$ there exists a decreasing sequence $\{l_k\}$ such that $l_k \to 0$ as $k \to \infty$ and*

$$\Pr\left\{\left|\frac{\zeta_n(k)}{1+k} - m\right| > l_k \textit{ for some integer } k \ge 0\right\} \le \frac{\varepsilon^n}{4}$$

for all $n = 1, 2, \dots$.

Proof. Let $\delta > 0$ be a constant such that (9) holds whenever $a \le \delta$. Let $\{\zeta_n^i(t)\}$ denote the i-th component of the process $\{\zeta_n(t)\}$. By Chebyshev's inequality, for $l \ge 0$ and $|\eta| \le \delta$ we have

$$\Pr\{\zeta_n^i(s) > l\} \le \exp\{nsq^i(\eta) - n\eta l\}, \tag{42}$$

where $q^i(\eta) = q(\theta^1, \dots, \theta^d)$ with $\theta^i = \eta$ and $\theta^j = 0$ for $j \ne i$. Recall that $\mathsf{E}\zeta(t) = mt$ for $t \ge 0$. Let m^i denote the i-th component of m. By the assumption (9) we have

$$\frac{\partial q^i}{\partial \eta}(0) = m^i \text{ and } \frac{\partial^2 q^i}{\partial \eta^2}(0) < \infty.$$

Consequently, there are constants $a^i > 0$ and $\delta_1 > 0$ such that

$$q^i(\eta) - \eta m^i \le a^i \eta^2$$

whenever $|\eta| \le \delta_1$. Then for any $l > 0$ we can find a constant $\eta = \eta(l) > 0$ such that

$$q^i(\eta) - \eta m^i - \eta l < 0. \tag{43}$$

By (42) and (43), for some constant $c_i^+(l) > 0$ we have

$$\Pr\{\zeta_n^i(s) > sl + m^i s\} \le \exp\{-nsc_i^+(l)\}. \tag{44}$$

In a similar way, we find the constant $c_i^-(l) > 0$ such that

$$\Pr\{\zeta_n^i(s) < -sl + m^i s)\} \le \exp\{-nsc_i^-(l)\}. \tag{45}$$

Summing up the inequalities (44) and (45) over $i = 1, \dots, d$ we see that for any $l > 0$ there exists a constant $c(l) > 0$ such that

$$\Pr\{|\zeta_n(s) - sm| > sl\} \le 2d \exp\{-nsc(l)\}$$

for all $s \ge 0$. Of course, we have $c(l) \to 0$ as $l \to 0$. Nevertheless we can choose a monotone sequence $\{l_k\}$ such that $kc(l_k) \ge \sqrt{k}$ and $l_k \to 0$ as $k \to \infty$. It follows that

$$\Pr\{|\zeta_n(k) - km| > kl_k\} \le 2d \exp\left\{-n\sqrt{k}\right\}.$$

Using the inequality (41) we can find sufficiently large $p = p(\varepsilon)$ so that

$$\sum_{k=p}^{\infty} \Pr\{|\zeta_n(k) - km| > kl_k\} \le \frac{\varepsilon^n}{8}.$$

This implies that

$$\Pr\{|\zeta_n(k) - km| \le kl_k \text{ for all integers } k \ge p\} \ge 1 - \frac{\varepsilon^n}{8},$$

and hence

$$\Pr\left\{\left|\frac{\zeta_n(k)}{1+k} - m\right| \le l_k + \frac{|m|}{1+k} \text{ for all integers } k \ge p\right\} \ge 1 - \frac{\varepsilon^n}{8}, \tag{46}$$

From Lemma 9 we derive that

$$\Pr\left\{\left|\frac{\zeta_n(k)}{1+k} - m\right| \le R \text{ for all integers } 0 \le k < p\right\} \ge 1 - \frac{\varepsilon^n}{8}. \tag{47}$$

for large enough constant $R = R(\varepsilon)$. Now the desired result follows from (46) and (47). □

Lemma 12 *For any $\varepsilon > 0$ there exists a decreasing sequence $\{d_k\}$ such that $d_k \to 0$ as $k \to \infty$ and*

$$\Pr\left\{\frac{|\zeta_n(k+1) - \zeta_n(k)|}{1+k} \ge d_k \text{ for some integer } k \ge 0\right\} \le \frac{\varepsilon^n}{4}. \tag{48}$$

for all $n = 1, 2, \ldots$.

Proof. Let $\eta \le a$ (see(9)). By Chebyshev's inequality, for $l > 0$ we have

$$\Pr\{|\zeta_n(k+1) - \zeta_n(k)| > l\} \le \exp\{-n\eta l + nq(\eta, \ldots, \eta)\}. \tag{49}$$

We take a constant $A > 0$ and let

$$d_k = \frac{1}{(k+1)\eta}\left[q(\eta, \ldots, \eta) + A + 2\ln(k+1)\right]. \tag{50}$$

It follows from (49) and (50) that

$$\sum_{k=0}^{\infty} \Pr\left\{\frac{|\zeta_n(k+1) - \zeta_n(k)|}{1+k} > d_k\right\} \le e^{-nA} \sum_{k=0}^{\infty} \frac{1}{(k+1)^2}.$$

Then for any $\varepsilon > 0$ we can choose $A = A(\varepsilon)$ so large that (48) holds for all $n = 1, 2, \ldots$. □

Lemma 13 *For any $\varepsilon > 0$ there exists a right continuous, nonnegative, decreasing function D_ε on $[0,\infty)$ such that $D_\varepsilon(t) \to 0$ as $t \to \infty$ and*

$$\Pr\left\{\left|\frac{\zeta_n(t)}{1+t} - m\right| > D_\varepsilon(t) \text{ for some } t \geq 0\right\} \leq \frac{\varepsilon^n}{2} \tag{51}$$

for all $n = 1, 2, \ldots$.

Proof. Let the sequences $\{l_k\}$ and $\{d_k\}$ be provided by Lemmas 11 and 12, and let $D_\varepsilon(k) = 2l_k + d_k + |m|/(k+1)$. Then $\{D_\varepsilon(k)\}$ is decreasing and $D_\varepsilon(k) \to 0$ as $k \to \infty$. Let $D_\varepsilon(t) = D_\varepsilon([t])$, where $[t]$ denotes the integer part of $t \geq 0$. Observe that for $k \leq s < k+1$ we have

$$\begin{aligned}\left|\frac{\zeta_n(s)}{1+s} - m\right| &\leq \left|\frac{\zeta_n(k)}{1+k} - m\right| + \frac{|\zeta_n(s) - \zeta_n(k)|}{1+s} + \frac{|\zeta_n(k)|(s-k)}{(1+k)(1+s)} \\ &\leq \left|\frac{\zeta_n(k)}{1+k} - m\right| + \frac{|\zeta_n(s) - \zeta_n(k)|}{1+k} + \frac{1}{1+s}\left|\frac{\zeta_n(k)}{1+k} - m\right| + \frac{|m|}{1+s} \\ &\leq 2\left|\frac{\zeta_n(k)}{1+k} - m\right| + \frac{|\zeta_n(s) - \zeta_n(k)|}{1+k} + \frac{|m|}{1+k}.\end{aligned}$$

Then (51) follows immediately from the Lemmas 11 and 12. □

References

[1] Shwartz, A. and Weiss, A. (1995). *Large Deviations for Performance Analysis: Queues, Communication and Computing.* Chapman and Hall, New York.

[2] Dembo, A. and Zeitouni, O. (1992). *Large Deviations Techniques and their Applications.* Jones and Bartlett Publ. Boston.

[3] Bertsimas, D., Paschalidis, I.Ch. and Tsitsiklis, J.N. (1998). On the large deviations behavior of acyclic networks of G/G/1 queues. *Ann. Appl. Probab.* **8** 1027–1069.

[4] Ramanan, K. and Dupuis, P. (1998). Large deviation properties of data streams that share a buffer. *Ann. Appl. Probab.* **8** 1070–1129.

[5] Dobrushin R.L., Pechersky E.A. (1994). Large deviations for tandem queuing systems. *J. of Applied Math. and Stoch. Analysis.* **7** 301–330.

[6] Aspandiiarov, S. and Pechersky, E.A. (1997). One large deviation problem for compound Poisson processes in queueing theory. *Markov Proc. Appl.* **3** 333–366 .

[7] Deuschel, J.-D. and Stroock, D.W. (1989). *Large Deviations.* Academic Press, London.

[8] Varadhan, S.R.S. (1984). *Large Deviations and Applications.* Society for Industrial and Applied Mathematics, Philadelphia.

[9] Gulinsky, O.V. and Veretennikov, A.Yu. (1993.) *Large Deviations for Discrete-Time Processes with Averaging.* VSP, Utrecht.

[10] Ignatyuk, A., Malyshev, V.A. and Scherbakov, V.V. (1994.) Boundary effects in large deviation problems. *Uspehi Matematicheskih Nauk* **49** 43–102 (in Russian).

[11] Puhalskii, A. (1995.) Large deviation analysis of the single server queue *Queueing Systems* **21** 5–66.

[12] Dobrushin, R.L., Pechersky, E.A. (1998). Large deviations for random processes with independent increments on infinite intervals. *Problemy peredachi informacii.* **34** 76–108 (in Russian)

[13] Liptser, R.Sh. and Pukhalskii, A.A. (1992). Limit theorems on large deviations for semimartingales. *Stochastics and Stochastics Reports* **38** 201–249.

[14] Lynch, J. and Sethuraman, J. (1987). Large deviations for processes with independent increments. *Ann. Prob.* **15** 610–627.

[15] Mogulskii, A.A. (1993). Large deviations for processes with independent increments. *Ann.Prob.* **21** 202–215.

[16] Pukhalskii, A.A. (1991). On functional principle of large deviations. in: *New Trends in Probability and Statistics.* eds. V.Sazonov and T.Shervashidze, VSP/Mokslas. 198–219.

[17] Rockafellar, R.T. (1970). *Convex Analysis.* Princton University Press, Princton, New Jersey.

[18] Dawson, D.A. and Gärtner, J. (1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308.