The Annals of Applied Probability
2000, Vol. 10, No. 3, 962–1001

LARGE DEVIATIONS OF JACKSON NETWORKS

BY IRINA IGNATIOUK-ROBERT

Université de Cergy-Pontoise

The problem of large deviations for a Jackson network is analyzed in detail. A new representation of the rate function is given and a simple procedure is proposed to get its closed form expression. The methods used rely on twisted distributions, localized processes, fluid limits and a careful analysis of some functions.

1. Introduction. We consider an open Jackson network with N queues. For $i = 1, \dots, N$, the arrivals at the ith queue are Poisson with parameter λ_i and the services delivered by the server are exponentially distributed with parameters μ_i. All the Poisson processes and the services are assumed to be independent. The routing matrix is denoted $(p_{ij};\, i, j = 1, \dots, N)$, p_{ij} is the probability that a customer goes to the jth queue when it has finished its service at queue i. The residual quantity

$$p_{i0} = 1 - \sum_{j=1}^{N} p_{ij}$$

is the probability that this customer definitively leaves the network. If the initial lengths of the queues are given by the vector $x = (x_i;\, i = 1, \dots, N)$, $X_i(t, x)$ denotes the length of the queue i at time t. The process $(X(t, x))$ is a continuous time Markov process on $\mathbb{Z}_+^N$ with $X(0, x) = x$. Its generator is given by

$$\mathscr{L} f(y) = \sum_{z \in \mathbb{Z}_+^N} q(z - y)(f(z) - f(y)), \qquad y \in \mathbb{Z}_+^N,$$

where

$$q(y) = \begin{cases} \lambda_i, & \text{if } y = \epsilon^i,\ i \in \{1, \dots, N\}, \\ \mu_i p_{i0}, & \text{if } y = -\epsilon^i,\ i \in \{1, \dots, N\}, \\ \mu_i p_{ij}, & \text{if } y = \epsilon^j - \epsilon^i,\ i, j \in \{1, \dots, N\}, \\ 0, & \text{otherwise,} \end{cases} \tag{1.1}$$

ϵ^i is the ith unit vector, $\epsilon^i_j = 0$ if $j \neq i$ and $\epsilon^i_j = 1$ otherwise. We set $p_{00} = 1$ and $p_{0i} = 0$ for $i \neq 0$, the matrix $(p_{ij};\, i, j = 0, \dots, N)$ is then stochastic.

ASSUMPTION A. We suppose that the matrix $(q(x - y);\ x, y \in \mathbb{Z}^N)$ is irreducible.

Received May 1999; revised December 1999.

AMS 1991 *subject classifications*. Primary 60F10; secondary 60J15, 60K35.

Key words and phrases. Large deviations, Jackson networks, twisted distributions, fluid limits, localized processes.

Assumption A means that the following conditions are satisfied jointly:

(i) The spectral radius of the matrix $(p_{ij};\, i, j = 1, \dots, N)$ is strictly less than unity.

(ii) For any $i = 1, \dots, N$, there exist $n \in \mathbb{N}$ and $j \in \{1, \dots, N\}$ such that $\lambda_j p_{ji}^{(n)} > 0$, where $p_{ij}^{(n)}$ is the n-time transition probabilities of a Markov chain with $N+1$ states associated to the stochastic matrix $(p_{ij};\, i, j = 0, \dots, N)$.

The spectral radius of the matrix $(p_{ij};\, i, j = 1, \dots, N)$ is strictly less than unity if and only if a customer leaves the network with probability 1; that is for any $i \in \{1, \dots, N\}$ there exists $n \in \mathbb{N}$, such that $p_{i0}^{(n)} > 0$.

We recall some of the well-known results concerning Jackson networks; see [20], for example. The traffic equations of the Jackson network are the following system of equations:

$$\nu_j = \lambda_j + \sum_{i=1}^{N} \nu_i p_{ij}, \qquad j = 1, \dots, N. \tag{1.2}$$

Under Assumption A, the system has a unique solution (ν_i) and the Markov process $(X(t, x))$ is ergodic if and only if

$$\nu_i < \mu_i \quad \text{for all } i = 1, \dots, N. \tag{1.3}$$

Under these conditions the stationary probabilities $(\pi(x);\, x \in \mathbb{Z}_+^N)$ of the Markov process $(X(x, t))$ are given by the product formula

$$\pi(x) = \prod_{i=1}^{N} \rho_i^{x_i}(1 - \rho_i), \qquad x \in \mathbb{Z}_+^N, \tag{1.4}$$

where $\rho_i = \nu_i/\mu_i$ for $i = 1, \dots, N$. When the network is at equilibrium, the components of the Markov process are independent.

A functional strong law of large numbers has been established by Chen and Mandelbaum in [6]; it shows that almost surely,

$$\lim_{n \to +\infty} \frac{1}{n} X(nt, [nx]) = Q(t, x),$$

uniformly on $t \in K$ for any compact set K. The process $(Q(t, x))$ is a deterministic process also called a fluid limit or fluid approximation of the Jackson network. Up to now there are different characterizations of the fluid limit $(Q(t, x))$. Chen and Mandelbaum [6] describe $(Q(t, x))$ in terms of the oblique reflection mapping, which is a solution of the associated Skorokhod problem (see [13]). Botvich and Zamyatin [5] characterize $(Q(t, x))$ in terms of the second vector field. They give an explicit expression of that by using the explicit form for stationary probabilities (1.4). We shall recall some of these results in Section 6.

A natural step after the analysis of the fluid limit is the study of the large deviation problem. Denote by $\mathscr{D}([0, T], \mathbb{R}_+^N)$ the set of cadlag functions from

$[0, T]$ to $\mathbb{R}_+^N$ endowed with Skorokhod topology. For $x \in \mathbb{R}_+^N$ and $t \in [0, T]$, the renormalized process $(Z_n(t, x))$ is defined by

$$Z_n(t, x) = \frac{1}{n} X(nt, [nx]).$$

Following the usual terminology, the Markov process is said to satisfy the sample path large deviation principle with the good rate function

$$I_{x,T}(\cdot) : \mathscr{D}([0, T], \mathbb{R}_+^N) \to \mathbb{R}_+,$$

iff for every $x \in \mathbb{R}_+^N$ and $T > 0$:

1. For any $c \in R_+$, the level set $\{\varphi : I_{x,T}(\varphi) \leq c\}$ is a compact subset of $\mathscr{D}([0, T], \mathbb{R}_+^N)$.
2. For every open set G of $\mathscr{D}([0, T], \mathbb{R}_+^N)$,
$$\liminf_{n\to+\infty} \frac{1}{n} \log P\big(Z_n(\cdot, x) \in G\big) \geq - \inf_{\varphi\in G} I_{x,T}(\varphi) .$$
3. For every closed set F of $\mathscr{D}([0, T], \mathbb{R}_+^N)$,
$$\limsup_{n\to+\infty} \frac{1}{n} \log P\big(Z_n(\cdot, x) \in F\big) \leq - \inf_{\varphi\in F} I_{x,T}(\varphi).$$

Dupuis and Ellis [10] proved the sample path large deviation principle with a good rate function for a large class of Markov processes describing various queueing systems satisfying the communication condition (see [10], Condition 3.4). This class of processes includes in particular open Jackson networks for which the communication condition holds whenever Assumption A is satisfied. However, this paper does not give an explicit representation for the rate function.

We consider the problem of the rate function identification, which is important in view of the possible applications. The main difficulty arises here from a discontinuity in the transition mechanism of the process. The Markov process $(X(t, x))$ describing the Jackson network is a process with discontinuous statistics; the dynamic is discontinuous at the boundary set

$$\big\{x \in \mathbb{Z}_+^N : x_i = 0, \text{ for some } 1 \leq i \leq N\big\},$$

that is, the ith component is decreased by 1 at rate μ_i only if it is positive. This discontinuity property is a serious difficulty for the large deviation analysis.

For tandem queues, the large deviation analysis and the problem of the rate function identification have been carried out by using the contraction principle [8, 29]. It is not usually possible to apply this method in more general situations and in particular for open Jackson networks. General results and explicit representations of the rate function were obtained for Markov processes with a discontinuity in the transition mechanism along a hyperplane, or more generally along a smooth $(n-1)$-dimensional interface in $\mathbb{R}^N$ (see [1, 3, 9, 17, 28]).

To identify the rate function for the Markov processes describing open Jackson networks, Dupuis, Ishii and Soner [12] proposed the method of viscosity solutions. It gives a variational expression of the rate function. Atar and Dupuis [2] extended these results for a more general class of the Markov processes when the associated Skorokhod problem has some regularity properties. They used a probabilistic method and simplified the representation of the rate function. However, their description is rather implicit; it does not give a simple procedure for its calculation.

We propose another probabilistic method for the large deviation analysis of Jackson networks. The method is based on a change of probability measure using an exponential martingale and on a careful analysis of the related fluid limits. We give an explicit form of the rate function and an algorithm to get its closed form expression.

The next section presents an overview of our results.

2. The main results. Articles [10] and [2] show that the Markov process $(X(t,x))$ satisfies the sample path large deviation principle with the rate function given by

$$I_{x,T}(\varphi) = \begin{cases} \int_0^T L(\varphi(t), \dot{\varphi}(t))\, dt, & \text{if } \varphi \text{ is absolutely continuous,} \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.1}$$

The function $L(x, v)$ is uniquely determined by the following limits:

$$\begin{aligned} &\frac{1}{\tau} \lim_{\delta\to 0} \lim_{\varepsilon\to 0} \limsup_{n\to+\infty} \frac{1}{n} \sup_{y:|y-nx|<\varepsilon n} \log \mathbb{P}\Bigl(\sup_{t\in[0,n\tau]} |X(t,y) - nx - vt| < \delta n\Bigr) \\ &\quad = \frac{1}{\tau} \lim_{\delta\to 0} \lim_{\varepsilon\to 0} \liminf_{n\to+\infty} \frac{1}{n} \inf_{y:|y-nx|<\varepsilon n} \log \mathbb{P}\Bigl(\sup_{t\in[0,n\tau]} |X(t,y) - nx - vt| < \delta n\Bigr), \end{aligned}$$

the constant $\tau > 0$ is supposed to satisfy $x + vt \in \mathbb{R}_+^N$ for all $t \in [0, \tau]$; otherwise $L(x, v) = +\infty$ if such a τ does not exist.

The function $L(\cdot,\cdot)$ completely determines the rate function; its explicit representation is the main goal of this paper. We introduce some definitions used throughout this paper.

DEFINITION 1. The function R is defined by

$$R(\alpha) = \sum_{y\in\mathbb{Z}^N} q(y)\bigl(e^{\langle\alpha, y\rangle} - 1\bigr), \qquad \alpha \in \mathbb{R}^N,$$

$\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^N$. If Λ is a subset of $\{1, \dots, N\}$, the quantity $l^\Lambda(v)$ is given by

$$l^\Lambda(v) = \sup_{\alpha\in\mathscr{B}_\Lambda} (\langle\alpha, v\rangle - R(\alpha)), \tag{2.2}$$

where $\mathscr{B}_\Lambda$ is the set

$$\mathscr{B}_\Lambda = \left\{\alpha \in \mathbb{R}^N : \alpha_i \le \log\left(\sum_{j=1}^N p_{ij} e^{\alpha_j} + p_{i0}\right) \text{ for all } i \notin \Lambda\right\}. \tag{2.3}$$

Using the definition (1.1), the function R can be expressed as

$$\begin{aligned} R(\alpha) = {} & \sum_{i=1}^N \mu_i\left(\sum_{j=1}^N p_{ij}\exp(\alpha_j - \alpha_i) + p_{i0}\exp(-\alpha_i) - 1\right) \\ & + \sum_{i=1}^N \lambda_i(\exp(\alpha_i) - 1). \end{aligned} \tag{2.4}$$

For a subset Λ of $\{1, \dots, N\}$, we denote $\Lambda^c = \{1, \dots, N\} \setminus \Lambda$; for $x \in \mathbb{R}_+^N$, we set

$$\Lambda(x) = \{i \in \{1, \dots, N\} : x_i > 0\},$$

and for $x \in \mathbb{R}^N$, $x_\Lambda = (x_i)_{i\in\Lambda}$.

The first result of the paper is the following theorem.

THEOREM 1. *For any $x \in \mathbb{R}_+^N$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c(x)} = 0$,*

$$L(x, v) = l^{\Lambda(x)}(v). \tag{2.5}$$

REMARK 2.1. The rate function $I_{x,T}(\cdot)$ is completely determined by the values $L(x, v)$ where $x \in \mathbb{R}_+^N$ and $v \in \mathbb{R}^N$ are such that $v_{\Lambda^c(x)} = 0$. Indeed, equation $v_{\Lambda^c(x)}=0$ is satisfied iff for all index i, $v_i = 0$ when $x_i = 0$. If $t \in (0, T)$ is such that $\varphi_i(t) = 0$, then for all $s \ne t$, we have $\varphi_i(t) \le \varphi_i(s)$ and hence, if $\dot\varphi_i(t)$ exists, it must be zero. Consequently, if φ is an absolutely continuous function in $\mathscr{D}([0, T], \mathbb{R}_+^N)$, then for all index i and for almost all $t \in [0, T]$, $\dot\varphi_i(t) = 0$ when $\varphi_i(t) = 0$ and therefore, the representation (2.1) of the rate function $I_{x,T}(\cdot)$ shows that only the values of $L(x, v)$ for x and v such that $v_{\Lambda^c(x)} = 0$ are really used.

Definition (2.2) of the function $l^\Lambda(v)$ is not easy to use in practice because the set $\mathscr{B}_\Lambda$ where the supremum (2.2) is taken is not convex. To get a more explicit form for $l^\Lambda(v)$ we identify the point $\alpha^v \in \mathscr{B}_\Lambda$ where the supremum (2.2) is achieved.

To prove the existence of the above α^v it is useful to consider the following system of equations:

$$\begin{cases} \beta_i = \alpha_i, & i \in \Lambda, \\ \beta_i = -\log\left(\sum_{j=1}^N p_{ij}\exp(\alpha_j - \alpha_i) + p_{i0}\exp(-\alpha_i)\right), & i \in \Lambda^c. \end{cases} \tag{2.6}$$

We prove that the unique solution $\alpha(\beta)$ of this system is a diffeomorphism from the convex set

$$\mathbb{R}_{\le 0}^{\Lambda, N} = \{\beta \in \mathbb{R}^N : \beta_i \le 0 \text{ for all } i \in \Lambda^c\}$$

to $\mathscr{B}_\Lambda$. Moreover, we show that the function $\beta \to R(\alpha(\beta))$ is strictly convex on $\mathbb{R}^{\Lambda, N}_{\le 0}$ and for $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, the level sets of the function $\beta \to R(\alpha(\beta)) - \langle \alpha, v\rangle$ are compact in $\mathbb{R}^{\Lambda, N}_{\le 0}$ (Proposition 8.1). In particular, for a given subset Λ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, this will prove the existence of a unique $\beta^v \in \mathbb{R}^{\Lambda, N}_{\le 0}$, and consequently a unique $\alpha^v = \alpha(\beta^v) \in \mathscr{B}_\Lambda$, such that

$$l^\Lambda(v) = \langle \beta^v, v\rangle - R(\alpha(\beta^v)) = \langle \alpha^v, v\rangle - R(\alpha^v);$$

β^v is the unique local minimum of the function $\beta \to R(\alpha(\beta)) - \langle \beta, v\rangle$ in $\mathbb{R}^{\Lambda, N}_{\le 0}$, and $\alpha^v = \alpha(\beta^v)$ is the unique local minimum of the function $\alpha \to R(\alpha) - \langle \alpha, v\rangle$ in $\mathscr{B}_\Lambda$.

The next step analyzes the location of α^v and gives an explicit form for $l^\Lambda(v)$. For a subset of indices Λ, the set $\mathscr{D}_\Lambda$ is defined as

$$\mathscr{D}_\Lambda = \left\{\alpha \in \mathbb{R}^N : \alpha_i = \log\left(\sum_{j=1}^N p_{ij} e^{\alpha_j} + p_{i0}\right), \ \forall\, i \in \Lambda^c\right\},$$

and $H^*_\Lambda(\cdot)$ by

$$H^*_\Lambda(v) = \sup_{\alpha \in \mathscr{D}_\Lambda} \{\langle \alpha, v\rangle - R(\alpha)\}.$$

We prove that for $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, $H^*_\Lambda(v)$ is the Fenchel–Legendre transform of the function $\alpha_\Lambda \to H_\Lambda(\alpha_\Lambda)$ given by

$$\begin{aligned} H_\Lambda(\alpha_\Lambda) = {} & \sum_{i \in \Lambda}\left(\nu_i - \sum_{j \in \Lambda} \nu_j m^\Lambda_{ji}\right)(e^{\alpha_i} - 1) \\ & + \sum_{i \in \Lambda} \mu_i \left(\sum_{j \in \Lambda} m^\Lambda_{ij} \exp(\alpha_j - \alpha_i) + m^\Lambda_{i0} \exp(-\alpha_i) - 1\right), \end{aligned}$$

where (ν_i) is the solution of the traffic equations (1.2) and

$$m^\Lambda_{ij} = p_{ij} + \sum_{k \ge 1} \sum_{j_1, \dots, j_k \in \Lambda^c} p_{ij_1} \cdots p_{j_k j}, \qquad i, j \in \{0, 1, \dots, N\}.$$

For $j \in \Lambda$, the value m^Λ_{ij} is the probability that the Markov chain associated with the stochastic matrix $(p_{ij};\ i, j = 0, \dots, N)$ reaches the point j without visiting the set Λ in a mean time given that it starts in i. For $j \in \Lambda^c$, m^Λ_{ij} is the mean time that the above Markov chain spends at the point j before hitting the set $\Lambda \cup \{0\}$.

The set $\mathscr{D}_\Lambda$ being a subset of $\mathscr{B}_\Lambda$, we have the following inequality:

$$H^*_\Lambda(v) = \sup_{\alpha \in \mathscr{D}_\Lambda} \{\langle \alpha, v\rangle - R(\alpha)\} \le \sup_{\alpha \in \mathscr{B}_\Lambda} \{\langle \alpha, v\rangle - R(\alpha)\} = l^\Lambda(v).$$

Let α^v be the point which achieves the maximum in $\mathscr{B}_\Lambda$. We give a necessary and sufficient condition for $\alpha^v \in \mathscr{D}_\Lambda$. When it is the case, we have clearly

$$l^\Lambda(v) = H^*_\Lambda(v)$$

and the quantity $l^{\Lambda}(v)$ is given therefore by an explicit function. When $\alpha^v \notin \mathscr{D}_\Lambda$, we prove that

$$l^{\Lambda}(v) = \min_{j \in \Lambda^c} l^{\Lambda \cup \{j\}}(v).$$

This gives a recursive procedure to calculate $l^{\Lambda}(v)$:

1. If $\alpha^v \in \mathscr{D}_\Lambda$, then $l^{\Lambda}(v) = H^*_\Lambda(v)$.
2. Otherwise, for each $j \in \Lambda^c$, calculate $l^{\Lambda \cup \{j\}}(v)$.

This algorithm terminates since for $\Lambda = \{1, \dots, N\}$, we have $\mathscr{B}_\Lambda = \mathscr{D}_\Lambda = \mathbb{R}^N$ and consequently,

$$l^{\{1,\dots,N\}}(v) = H^*_{\{1,\dots,N\}}(v) = R^*(v),$$

where $R^*(\cdot)$ is the Fenchel–Legendre transform of the function $R(\cdot)$. We give an example of that procedure in Section 3.

The following theorem summarizes the above results.

THEOREM 2. *For $\Lambda \subseteq \{1, \dots, N\}$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$:*

(i) *There is a unique point $\tilde{\alpha}^v$ of $\mathscr{D}_\Lambda$ such that*

$$H^*_\Lambda(v) = \langle \tilde{\alpha}^v, v\rangle - R(\tilde{\alpha}^v), \tag{2.7}$$

$\tilde{\alpha}^v_\Lambda$ is a unique solution of the system $\nabla\, H_\Lambda(\alpha_\Lambda) = v_\Lambda$ and

$$\tilde{\alpha}^v_i = \log\left(\sum_{j\in\Lambda} m^\Lambda_{ij} \exp(\tilde{\alpha}^v_j) + m^\Lambda_{i0}\right), \qquad i \in \Lambda^c.$$

(ii) *The equality $l^{\Lambda}(v) = H^*_\Lambda(v)$ holds if and only if*

$$\nu^\Lambda_j(\tilde{\alpha}^v) = \left(\nu_j + \sum_{i\in\Lambda} m^\Lambda_{ij}\left(\mu_i \exp(-\tilde{\alpha}^v_i) - \nu_i\right)\right)\exp(\tilde{\alpha}^v_j) \le \mu_j \text{ for all } j \in \Lambda^c. \tag{2.8}$$

(iii) *Otherwise,*

$$l^{\Lambda}(v) = \min_{j\in\widetilde{\Lambda}} l^{\Lambda\cup\{j\}}(v),$$

where

$$\widetilde{\Lambda} = \left\{ j \in \Lambda^c \colon \nu^\Lambda_j(\tilde{\alpha}^v) > \mu_i \right\}.$$

The above theorem together with Theorem 1 implies in particular that for $x = 0$ and $v = 0$ ($\Leftrightarrow \Lambda(x) = \varnothing$):

1. $L(0,0) = 0$ if and only if $\nu_i \le \mu_i$ for all $i = 1, \dots, N$.
2. Otherwise,

$$L(0,0) = \min_{i\in\widetilde{\Lambda}} \left\{ l^{\{i\}}(0) \right\} = \min_{i\in\widetilde{\Lambda}} \left\{ L(\varepsilon^i, 0) \right\} > 0,$$

where (ν_i) is the unique solution of the traffic equations (1.2),

$$\widetilde{\Lambda} = \{i : \nu_i > \mu_i\}$$

and $\epsilon^i \in \mathbb{R}_+^N$ is the ith unit vector that is, $\varepsilon_j^i = 1$ if $i = j$ and $\varepsilon_j^i = 0$ otherwise.

Our paper is organized as follows. To prove Theorem 1 we verify the local lower large deviation bound,

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\liminf_{n\to+\infty}\frac{1}{n}\inf_{y:|y-nx|<n\varepsilon}\log\mathbb{P}\left(\sup_{t\in[0,n\tau]}|X(t,y)-nx-vt|<\delta n\right) \geq -\tau l^{\Lambda(x)(v)}, \tag{2.9}$$

and the local upper large deviation bound,

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\sup_{y:|y-nx|<n\varepsilon}\log\mathbb{P}\left(\sup_{t\in[0,n\tau]}|X(t,y)-nx-vt|<\delta n\right) \leq -\tau l^{\Lambda(x)}(v). \tag{2.10}$$

To simplify the left-hand side of the above bounds it is convenient to rewrite them in terms of local models. A local model relative to $x \in \mathbb{R}_+^N$ is the simplest Markov process with the same large deviation behavior in a neighborhood of $x \in \mathbb{R}_+^N$ as the original process $(X(t,x))$. We introduce the local models and we rewrite the above bounds in terms of that in Section 5. To prove the lower bound we use the results of [5]. We recall these results in Section 6. Section 7 is devoted to the exponential change of measure and the associated twisted Markov processes. We show that a twisted Markov process is also an open Jackson network, and we describe its fluid approximation. In Section 8 we establish some properties of the functions $l^\Lambda(\cdot)$ being used in the proof of the local lower large deviation bound. In particular, we relate the functions $l^\Lambda(\cdot)$ with the fluid approximation of twisted Markov processes. In Section 9 the proof of the local large deviation bounds is completed and Section 10 is devoted to the proof of Theorem 2.

Two examples illustrate our results. We begin with an example of a Jackson network with two nodes in Section 3, which shows that our method is effective, that is, one can get explicit expressions for the rate function with our results. In Section 4 we show how the explicit expression of the rate function can be used to estimate the mean time until the total number of customers in an ergodic Jackson network reaches the level n, given that the process starts in 0.

3. Example: a Jackson network with two nodes. In this section, the Markov process $(X(t,x))$ describes an open Jackson network with two nodes $(N = 2)$. See Figure 1. Its generator is given by

$$\begin{aligned}\mathscr{L}f(x) = {}&\lambda_1\big(f(x+\epsilon_1)-f(x)\big)+\mu_1\big(p_{12}f(x-\epsilon_1+\epsilon_2)+p_{10}f(x-\epsilon_1)-f(x)\big)\\ &+\lambda_2\big(f(x+\epsilon_2)-f(x)\big)+\mu_2\big(p_{21}f(x-\epsilon_2+\epsilon_1)+p_{20}f(x-\epsilon_2)-f(x)\big).\end{aligned}$$

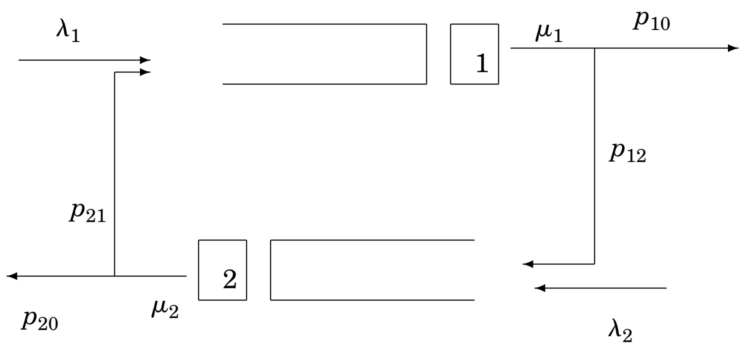

FIG. 1. *A Jackson network with two nodes.*

Assumption A is equivalent to:

1. $p_{12}p_{21} < 1$;
2. $\lambda_1 + \lambda_2 p_{21} > 0$ and $\lambda_2 + \lambda_1 p_{12} > 0$.

The solution of the traffic equations

$$\nu_1 = \lambda_1 + \nu_2 p_{21},$$
$$\nu_2 = \lambda_2 + \nu_1 p_{12}$$

is given by

$$\nu_1 = \frac{\lambda_1 + \nu_2 p_{21}}{1 - p_{12}p_{21}}, \qquad \nu_2 = \frac{\lambda_2 + \nu_1 p_{12}}{1 - p_{12}p_{21}}$$

and the function $R(\alpha) = R(\alpha_1, \alpha_2)$ is

$$\begin{aligned} R(\alpha_1, \alpha_2) = {} & \lambda_1(\exp(\alpha_1) - 1) + \mu_1(p_{12}\exp(\alpha_2 - \alpha_1) + p_{10}\exp(-\alpha_1) - 1) \\ & + \lambda_2(\exp(\alpha_2) - 1) + \mu_2(p_{21}\exp(\alpha_1 - \alpha_2) + p_{20}\exp(-\alpha_2) - 1). \end{aligned}$$

To identify the rate function (2.1) here, we have to calculate $L(x, v)$:

1. For $x = (x_1, x_2)$ with $x_1, x_2 > 0$ and arbitrary $v = (v_1, v_2) \in \mathbb{R}^2$
2. For $x = (0, x_2)$ with $x_2 > 0$ and $v = (0, v_2)$, $v_2 \in \mathbb{R}$
3. For $x = (x_1, 0)$ with $x_1 > 0$ and $v = (v_1, 0)$, $v_1 \in \mathbb{R}$
4. And for $x = (0, 0)$, $v = (0, 0)$.

The first case is trivial since in the interior of the domain the process is equivalent to a homogeneous random walk, and consequently,

$$L(x, v) = R^*(v),$$

where $R^*(\cdot)$ is the Fenchel–Legendre transform of the function $R(\cdot)$ (see [7] for example). To calculate $L(x, v)$ for the second case, the results of [3, 9, 17] can be used, because the large deviation behavior of our process in a neighborhood of x does not depend on the boundary $\{z\colon z_2 = 0\}$ if $x = (0, x_2)$ with $x_2 > 0$ (see Section 5 for more details). The third case is similar.

The last case is easy if the Markov process $(X(t, y))$ is ergodic (i.e., iff $\nu_i < \mu_i$ for all $i = 1, \dots, N$). In this case we obviously have $L(0, 0) = 0$.

The most difficult case is when $x = (0, 0)$, $v = (0, 0)$ and the process $(X(t, y))$ is not ergodic; in this case the influence of the two boundaries occurs. The following statement is a direct consequence of the second and third part of Theorem 2.

PROPOSITION 3.1. *$L(0, 0) = 0$ if and only if $\nu_1 \le \mu_1$ and $\nu_2 \le \mu_2$. Otherwise,*

$$L(0,0) = l^{\{1\}}(0) \quad \text{if } \nu_1 \ge \mu_1 \text{ and } \nu_2 \le \mu_2,$$
$$L(0,0) = l^{\{2\}}(0) \quad \text{if } \nu_2 \ge \mu_2 \text{ and } \nu_1 \le \mu_1$$

and

$$L(0,0) = \min\{l^{\{1\}}(0), l^{\{2\}}(0)\} \quad \text{if } \nu_1 \ge \mu_1 \text{ and } \nu_2 \ge \mu_2,$$

where

$$l^{\{1\}}(0) = -\inf\left\{R(\alpha) \mid \alpha = (\alpha_1, \alpha_2) : \alpha_2 \le \log(p_{21}e^{\alpha_1} + p_{20})\right\}$$

and

$$l^{\{2\}}(0) = -\inf\left\{R(\alpha) \mid \alpha = (\alpha_1, \alpha_2) : \alpha_1 \le \log(p_{12}e^{\alpha_2} + p_{10})\right\}.$$

Now we identify the values of $l^{\{1\}}(0)$ and $l^{\{2\}}(0)$. For this, notice that (α_1, α_2) belongs to $\mathscr{D}_{\{1\}}$ if and only if $\alpha_2 = \log(p_{21}e^{\alpha_1} + p_{20}))$, and if it is the case then

$$\begin{aligned} R(\alpha_1, \alpha_2) &= H_{\{1\}}(\alpha_1) \\ &= (\lambda_1 + \lambda_2 p_{21})(e^{\alpha_1} - 1) + \mu_1(1 - p_{12}p_{21})(e^{-\alpha_1} - 1). \end{aligned}$$

It is clear that the minimum of the function $H_\Lambda(\alpha_1)$ is achieved in the point

$$\tilde{\alpha}_1 = \frac{1}{2}\log\left(\frac{\mu_1(1 - p_{12}p_{21})}{\lambda_1 + \lambda_2 p_{21}}\right) = \frac{1}{2}\log\left(\frac{\mu_1}{\nu_1}\right),$$

and hence,

$$\begin{aligned} \inf_{\alpha_1} H_{\{1\}}(\alpha_1) &= H_{\{1\}}(\tilde{\alpha}_1) \\ &= -(1 - p_{12}p_{21})(\sqrt{\mu_1} - \sqrt{\nu_1})^2. \end{aligned}$$

Using Theorem 2 applied for the case where $\Lambda = \{1\}$, we get the proposition.

PROPOSITION 3.2.

$$l^{\{1\}}(0) = (1 - p_{12}p_{21})(\sqrt{\mu_1} - \sqrt{\nu_1})^2$$

if

$$\lambda_2 + \lambda_2 p_{21}\biggl(\sqrt{\frac{\mu_1}{\nu_1}} - 1\biggr) + \mu_1 p_{12} + \mu_1 p_{12} p_{20}\biggl(\sqrt{\frac{\nu_1}{\mu_1}} - 1\biggr) \le \mu_2$$

and

$$l^{\{1\}}(0) = -\min_{(\alpha_1,\alpha_2)\in\mathbb{R}^2} R(\alpha_1, \alpha_2)$$

otherwise.

The expression of $l^{\{2\}}(0)$ is quite similar to that of $l^{\{1\}}(0)$.

PROPOSITION 3.3.

$$l^{\{2\}}(0) = (1 - p_{12}p_{21})(\sqrt{\mu_2} - \sqrt{\nu_2})^2$$

if

$$\lambda_1 + \lambda_1 p_{12}\biggl(\sqrt{\frac{\mu_2}{\nu_2}} - 1\biggr) + \mu_2 p_{21} + \mu_2 p_{21} p_{10}\biggl(\sqrt{\frac{\nu_2}{\mu_2}} - 1\biggr) \le \mu_1$$

and

$$l^{\{2\}}(0) = -\min_{(\alpha_1,\alpha_2)\in\mathbb{R}^2} R(\alpha_1, \alpha_2)$$

otherwise.

Notice finally that

$$\min_{(\alpha_1,\alpha_2)\in\mathbb{R}^2} R(\alpha_1, \alpha_2) = R(\alpha_1^0, \alpha_2^0),$$

where (α_1^0, α_2^0) is a unique solution of the system

$$\lambda_1 \exp(\alpha_1) + \mu_2 p_{21} \exp(\alpha_1 - \alpha_2) = \mu_1(p_{12} \exp(\alpha_2 - \alpha_1) + p_{10} \exp(-\alpha_1)),$$
$$\lambda_2 \exp(\alpha_2) + \mu_1 p_{12} \exp(\alpha_2 - \alpha_1) = \mu_2(p_{21} \exp(\alpha_1 - \alpha_2) + p_{20} \exp(-\alpha_2)).$$

4. An application to a problem of exit time. Let $(X(t,x))$ describe an ergodic open Jackson network with N nodes for which Assumption A holds. We suppose therefore, that

$$\frac{\nu_i}{\mu_i} < 1 \quad \text{for all } i = 1, \dots, N,$$

where (ν_i) is the solution of the traffic equations (1.2). In this section we apply Theorem 2 to prove the following proposition.

PROPOSITION 4.1. *If $\mathscr{T}_n$ is the first time when the total number of customers of an initially empty ergodic Jackson network is greater than n,*

$$\mathscr{T}_n = \inf\left\{t : \sum_{i=1}^{N} X_i(t,0) \geq n\right\},$$

then

$$\lim_{n\to+\infty} \frac{1}{n} \log \mathbb{E}\mathscr{T}_n = \min_{1\leq i\leq N} (\log \mu_i - \log \nu_i). \tag{4.1}$$

The classical approach in this domain is given by Freidlin–Wentzel method [16] (other approaches can be found in [21] and [28], e.g.). It consists in getting the explicit expression of the rate function and finding the optimal path, that is, the path that drives the process $X(nt,0)/n$ out of the domain $\{y \in \mathbb{R}_+^N\colon \sum_i y_i \leq 1\}$ and minimizes the rate function. Since the explicit expression of the rate function for Jackson networks was not known, Frater, Lennon and Anderson [15] proved (4.1) using a heuristic approach proposed by Borovkov, Ruget [26] and others for a $GI/GI/1$ queue (see also [23]). With Theorem 2 we are able to establish rigorously the relation (4.1).

PROOF OF PROPOSITION 4.1. To prove (4.1) we have to verify the upper bound,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\mathscr{T}_n \leq \min_{1\leq i\leq N} (\log \mu_i - \log \nu_i) \tag{4.2}$$

and the lower bound

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\mathscr{T}_n \geq \min_{1\leq i\leq N} (\log \mu_i - \log \nu_i). \tag{4.3}$$

To prove the upper bound we use an idea from [15] which shows that only the dominating queue is important (we do not assume however, that this dominating queue is unique). This means that the optimal path driving our process out of the domain $\{x\colon \sum_i x_i \leq 1\}$ should be linear and that it should pass on the boundary $\{x \in \mathbb{Z}_+^N\colon x_i = 0 \text{ for all } i \neq i_0\}$, where i_0 achieves the maximum $\max_i(\nu_i/\mu_i)$. Using this idea together with Theorem 2, we will find a linear path $\varphi\colon [0,T] \to \mathbb{R}_+^N$ starting in 0 and going out of the domain $\{y \in \mathbb{R}_+^N\colon \sum_i y_i \leq 1\}$, for which $I_{0,T}(\varphi) = \min_i(\log \mu_i - \log \nu_i)$. This will give the upper bound (4.2). The lower bound will be obtained from the explicit form of the stationary probabilities (1.4) and therefore we do not have to prove that this path is really optimal.

We start with the proof of the upper bound. Without any restriction of generality we will assume that

$$\max_i(\nu_i/\mu_i) = \nu_1/\mu_1. \tag{4.4}$$

For $\Lambda = \{1\}$, the function $H_\Lambda(\cdot)$ is given by

$$H_{\{1\}}(\alpha_1) = m_{10}^{\{1\}}\Big(\nu_1(e^{\alpha_1} - 1) + \mu_i(e^{-\alpha_1} - 1)\Big)$$

and the vector $(\nu_j^{\{1\}}(\alpha);\ j \neq 1)$ is defined by

$$\nu_j^{\{1\}}(\alpha) = \Big(\nu_j + m_{1j}^{\{1\}}(\mu_1 e^{-\alpha_1} - \nu_1)\Big)e^{\alpha_j}, \qquad j = 2, \dots, N.$$

Theorem 2 shows that for $v = (v_1, 0, \dots, 0)$, the identity

$$l^{\{1\}}(v) = H_{\{1\}}^*(v)$$

is verified if and only if for any $j \neq 1$,

$$\Big(\nu_j + m_{1j}^{\{1\}}(\mu_1 e^{-\tilde{\alpha}_1} - \nu_1)\Big)\Big(m_{j1}^{\{1\}} e^{\tilde{\alpha}_1} + m_{j0}^{\{1\}}\Big) \leq \mu_j, \tag{4.5}$$

where $\tilde{\alpha}_1$ is the unique solution of the equation

$$\frac{d}{d\alpha_1} H_{\{1\}}(\alpha_1) = m_{10}^{\{1\}}(\nu_1 e^{\alpha_1} - \mu_1 e^{-\alpha_1}) = v_1.$$

Thus, for $v_1 = m_{10}^{\{1\}}(\mu_1 - \nu_1)$:

(i) The solution of the above equation is $\tilde{\alpha}_1 = \log \mu_1 - \log \nu_1$.
(ii) $H_{\{1\}}^*(v) = v_1(\log \mu_1 - \log \nu_1)$ because clearly, $H_{\{1\}}(\tilde{\alpha}_1) = 0$.
(iii) The inequalities (4.5) are verified if $\nu_1/\mu_1 \geq \nu_i/\mu_i$ for all $i = 2, \dots, N$ and hence, $l^{\{1\}}(v) = v_1(\log \mu_1 - \log \nu_1)$ whenever (4.4) holds.

Define the vector $v = (v_1, 0, \dots, 0) \in \mathbb{R}^N$ by setting $v_1 = m_{10}^{\{1\}}(\mu_1 - \nu_1)$ and let $T = 1/v_1$. Then the linear path $\varphi(t) = vt$, $t \in [0, T]$ satisfies

$$I_{0,T}(\varphi) = Tl^{\{1\}}(v) = \log \mu_1 - \log \nu_1. \tag{4.6}$$

We are ready to prove the upper bound (4.2). It is known (see, e.g., [23]) that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\mathcal{T}_n = -\lim_{n\to\infty} \frac{1}{n} \log a(0, \Delta_n),$$

where $a(0, \Delta_n)$ is the probability that the embedded discrete-time Markov chain (X_k) starting from 0 reaches the set $\Delta_n = \{x \in \mathbb{Z}_+^N : \sum_{i=1}^N x_i = n\}$ before returning to the state 0.

For $0 < \varepsilon < 1$, $0 < \delta < \varepsilon/2$, and $n > 1/\varepsilon$, the Markov process $X(x, t)$ starting from $x = ([n\varepsilon], 0, \dots, 0) \in \mathbb{R}^N$ reaches the set Δ_n before the time nT without visiting the state 0 in the mean time whenever the event

$$\sup_{t\in[0, nT]} |X(x, t) - x - vt| < \delta n \tag{4.7}$$

occurs. Consequently, the probability that the embedded Markov chain starting from $x = ([n\varepsilon], 0, \dots, 0)$ reaches the set Δ_n before returning to the state 0 is greater than the probability of the event (4.7).

Furthermore, Assumption A implies that the embedded Markov chain starting from 0 reaches the state $x = ([n\varepsilon], 0, \dots, 0)$ before hitting the set Δ_n and before returning to 0 with probability greater than $\gamma^{[n\varepsilon]}$ where

$$\gamma = \min_{y\neq 0} q(y)\bigg(\sum_{y\neq 0} q(y)\bigg)^{-1},$$

and hence,

$$a(0,\Delta_n) \geq \gamma^{[n\varepsilon]} \times \mathbb{P}\left\{\sup_{t\in[0,\,nT]} |X(x,t) - x - vt| < \delta n\right\}$$

for all $0 < \varepsilon < 1$, $0 < \delta < \varepsilon/2$ and $n > 1/\varepsilon$.

Letting now $n \to \infty$, $\delta \to 0$ and then $\varepsilon \to 0$ and using then (4.6) together with (2.9), we get the upper bound,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\mathscr{T}_n = -\liminf_{n\to\infty} \frac{1}{n} \log a(0,\Delta_n) \leq Tl^{\{1\}}(v) = \log \mu_1 - \log \nu_1.$$

The lower bound (4.3) easily follows from the explicit form of the stationary probabilities (1.4). Indeed, the embedded Markov chain (X_k) has the same stationary probabilities as the original Markov process $X(t,0)$ and

$$\begin{aligned}\binom{N-1}{n+N}\left(\frac{\nu_1}{\mu_1}\right)^n &\geq \frac{1}{\pi(0)} \sum_{x\in\Delta_n} \pi(x)\\ &= \mathbb{E}\left(\sum_{k=0}^{\tau_0} \mathbb{1}_{\{X_k\in\Delta_n\}}\right) \qquad (4.8)\\ &\geq \mathbb{P}\big(X_k \in \Delta_n \text{ for some } k \leq \tau_0\big) = a(0,\Delta_n),\end{aligned}$$

where τ_0 is the first time when the Markov chain X_n returns to 0 starting from 0. The first relation holds here because for each $x \in \Delta_n$ we have

$$\pi(x)/\pi(0) \leq (\nu_1/\mu_1)^n$$

when (4.4) holds and the cardinality of the set $\Delta_n = \{x \in \mathbb{Z}_+^N : \sum_{i=1}^N x_i = n\}$ is equal to the number of choices of $N-1$ elements between $n+N$ elements. The second relation (4.8) is the classical representation of the stationary probability of a Markov chain (see, e.g., [14]). □

5. Localized Markov process. In this section we simplify the left-hand side of the bounds (2.9) and (2.10). We rewrite them in terms of the local models.

The local model (localized Markov process) relative to $x \in \mathbb{R}_+^N$ is defined as follows.

Given $x \in \mathbb{R}_+^N$ and $\Lambda = \Lambda(x)$, let $(X^\Lambda(t,y))_{t\geq 0}$ be a Markov process on

$$\mathbb{Z}_+^{\Lambda,N} = \{z \in \mathbb{Z}^N : z_i \geq 0 \text{ for all } i \in \Lambda^c\}$$

with initial state $y \in \mathbb{Z}_+^{\Lambda,N}$ and generator

$$\mathscr{L}^\Lambda f(x) = \sum_{z\in\mathbb{Z}_+^{\Lambda,N}} q(z-x)(f(z)-f(x)), \qquad x \in \mathbb{Z}_+^{\Lambda,N}, \tag{5.1}$$

where the transition intensities $q(z)$, $z \in \mathbb{Z}^N$, are defined by (1.1). The above Markov process describes a modified open Jackson network with the same parameters as the original Markov process $(X(t, y))$, but without the boundary conditions on the nodes $i \in \Lambda$. The queue lengths at the nodes $i \in \Lambda$ may be negative. We say that this Markov process describes an open Jackson network where the nodes $i \in \Lambda$ are free.

To identify the function $L(x, v)$, it is sufficient to verify the local bounds (2.9) and (2.10) for $\tau > 0$ small enough, so that for any $i \in \Lambda(x)$,

$$x_i + v_i t > 0 \quad \text{for all} \quad t \in [0, \tau].$$

In this case, for all $\delta > \varepsilon > 0$ small enough and for every $y \in \mathbb{Z}_+^N$ such that $|y - nx| < \varepsilon n$, the probability

$$\mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X(t, y) - nx - vt| < \delta n \Big)$$

does not depend on the transition intensities on the boundary

$$\bigcup_{i \in \Lambda(x)} \{z : z_i = 0\}$$

and consequently,

$$\begin{aligned} &\mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X(t, y) - nx - vt| < \delta n \Big) \\ &\qquad = \mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X^{\Lambda(x)}(t, y) - nx - vt| < \delta n \Big). \end{aligned} \tag{5.2}$$

Thus, the new Markov process $(X^{\Lambda}(t, y))$ has the same large deviation behavior in a neighborhood of x as the original process $(X(t, y))$ and hence, to prove the local bounds (2.9) and (2.10) for the original process it is sufficient to verify them for $(X^{\Lambda}(t, y))$.

It is convenient, moreover, to rewrite the left-hand side of the local bounds (2.9) and (2.10) as follows.

PROPOSITION 5.1. *Given $x \in \mathbb{R}_+^N$, consider $\Lambda = \Lambda(x)$, $v \in \mathbb{R}^N$ and $\tau > 0$ such that $v_{\Lambda^c} = 0$ and $x_i + \tau v_i > 0$ for all $i \in \Lambda$; then*

$$\begin{aligned} &\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \liminf_{n \to +\infty} \inf_{|y - nx| < \varepsilon n} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X(t, y) - nx - vt| < \delta n \Big) \\ &\qquad = \lim_{\delta \to 0} \liminf_{n \to +\infty} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X^{\Lambda}(t, 0) - vt| < \delta n \Big) \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} &\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \limsup_{n \to +\infty} \sup_{|y - nx| < \varepsilon n} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X(t, y) - nx - vt| < \delta n \Big) \\ &\qquad = \lim_{\delta \to 0} \limsup_{n \to +\infty} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t \in [0, n\tau]} |X^{\Lambda}(t, 0) - vt| < \delta n \Big). \end{aligned} \tag{5.4}$$

This proposition is a direct consequence of Proposition 3.7 in [10].

Thus, to verify the local large deviation bounds (2.9) and (2.10) it is sufficient to prove the lower bound,

$$\lim_{\delta\to 0}\liminf_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0)-vt|<\delta n\Big)\geq -\tau l^\Lambda(v)$$

and the upper bound,

$$\lim_{\delta\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0)-vt|<\delta n\Big)\leq -\tau l^\Lambda(v)$$

for all $\Lambda\subseteq\{1,\dots,N\}$, $\tau\geq 0$ and $v\in\mathbb{R}^N$ such that $v_{\Lambda^c}=0$.

The proof of the upper bound is quite simple (see Section 9). To prove the lower bound we shall use the following proposition.

PROPOSITION 5.2. *Let $\Lambda\subseteq\Lambda'\subseteq\{1,\dots,N\}$ and $v\in\mathbb{R}^N$ such that $v_{\Lambda^c}=0$. Then*

$$\begin{aligned}&\lim_{\delta\to 0}\liminf_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0)-vt|<\delta n\Big)\\ &\qquad\geq\lim_{\delta\to 0}\liminf_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^{\Lambda'}(t,0)-vt|<\delta n\Big).\end{aligned}\tag{5.5}$$

PROOF. Given $\varepsilon>0$, define $x^{n\varepsilon}\in\mathbb{Z}_+^{\Lambda,N}$ by setting

$$x_i^{n\varepsilon}=\begin{cases}[n\varepsilon v_i], & \text{if } i\in\Lambda,\\ [n\varepsilon], & \text{if } i\in\Lambda'\setminus\Lambda,\\ 0, & \text{otherwise}\end{cases}$$

and consider the trajectories of the Markov process $X^\Lambda(t,0)$ which first go from 0 to $x^{n\varepsilon}$ in time $t_\varepsilon=n\varepsilon$ without leaving the set $\{z:|z|<1/2\delta n\}$ and then follow the path $\psi^{n\varepsilon}(t)=x^{n\varepsilon}+vt$ so that

$$\sup_{t\in[0,n(\tau-\varepsilon)]}|X^\Lambda(t+n\varepsilon,0)-x^{n\varepsilon}-vt|<\delta' n.$$

It is clear that for every $\delta>0$, there exists $\varepsilon_0>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$ and for all $\delta'\in(0,\delta-N\varepsilon)$,

$$\begin{aligned}&\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0)-vt|<\delta n\Big)\\ &\qquad\geq\mathbb{P}\Big(\sup_{t\in[0,n\varepsilon]}|X^\Lambda(t,0)|<\tfrac{1}{2}\delta n\text{ and }X^\Lambda(n\varepsilon,0)=x^{n\varepsilon}\Big)\\ &\qquad\quad\times\mathbb{P}\Big(\sup_{t\in[0,n(\tau-\varepsilon)]}|X^\Lambda(t,x^{n\varepsilon})-x^{n\varepsilon}-vt|<\delta' n\Big)\end{aligned}\tag{5.6}$$

if $n>N/(\delta-N\varepsilon-\delta')$.

Choose now a sequence $y_0 = 0, \dots, y_{m-1}, y_m = x^{n\varepsilon}$ such that $m \le N\varepsilon n$ and for all $k = 0, \dots, m$,

$$q(y_k - y_{k-1}) > 0 \quad \text{and} \quad |y_k| < \tfrac{1}{2}\delta n$$

(the above sequence exists because of Assumption A). Then the included discrete time Markov chain relative to $X^\Lambda(t, 0)$ goes from 0 to $x^{n\varepsilon}$ in time m without leaving the set $\{z: |z| < \delta n/2\}$ with probability $\ge \tilde{\gamma}^m \ge \tilde{\gamma}^{N\varepsilon n}$ where

$$\tilde{\gamma} = \frac{\min_{y\neq 0} q(y)}{\sum_{y\neq 0} q(y)}$$

and hence

$$\mathbb{P}\Big(\sup_{t\in[0,\,n\varepsilon]} |X^\Lambda(t, 0)| < \tfrac{1}{2}\delta n \text{ and } X^\Lambda(n\varepsilon, 0) = x^{n\varepsilon}\Big) \ge \gamma^{n\varepsilon}$$

with $\gamma = N^{-1}\tilde{\gamma}^N \exp(-\sum_{y\neq 0} q(y))$. Using therefore, relation (5.6) together with the above inequality, we get

$$\begin{aligned} &\mathbb{P}\Big(\sup_{t\in[0,\,n\tau]} |X^\Lambda(t, 0) - vt| < \delta n\Big) \\ &\qquad \ge \gamma^{n\varepsilon} \times \mathbb{P}\Big(\sup_{t\in[0,\,n(\tau-\varepsilon)]} |X^\Lambda(t, x^{n\varepsilon}) - x^{n\varepsilon} - vt| < \delta' n\Big) \\ &\qquad \ge \gamma^{n\varepsilon} \times \mathbb{P}\Big(\sup_{t\in[0,\,n\tau]} |X^\Lambda(t, x^{n\varepsilon}) - x^{n\varepsilon} - vt| < \delta' n\Big). \end{aligned} \tag{5.7}$$

Notice finally that for $0 < \delta < \varepsilon$ and $n > 1/(\varepsilon - \delta')$, the probability in the right-hand side of (5.7) does not depend on the transition intensities of the Markov process $X^\Lambda(t, x^{n\varepsilon})$ on the boundary

$$\bigcup_{i\in\Lambda'\setminus\Lambda} \{z: z_i = 0\}$$

because the relation

$$|X^\Lambda(t, x^{n\varepsilon}) - x^{n\varepsilon} - vt| < \delta' n$$

implies

$$X_i^\Lambda(t, x^{n\varepsilon}) \ge n(\varepsilon - \delta') - 1 > 0 \quad \text{for all } i \in \Lambda' \setminus \Lambda,$$

and therefore, for $0 < \delta < \varepsilon$ and $n > 1/(\varepsilon - \delta')$,

$$\begin{aligned} &\mathbb{P}\Big(\sup_{t\in[0,\,n\tau]} |X^\Lambda(t, x^{n\varepsilon}) - x^{n\varepsilon} - vt| < \delta' n\Big) \\ &\qquad = \mathbb{P}\Big(\sup_{t\in[0,\,n\tau]} |X^{\Lambda'}(t, x^{n\varepsilon}) - x^{n\varepsilon} - vt| < \delta' n\Big) \\ &\qquad = \mathbb{P}\Big(\sup_{t\in[0,\,n\tau]} |X^{\Lambda'}(t, 0) - vt| < \delta' n\Big), \end{aligned}$$

where the last equality holds because the generator of the Markov process $X^{\Lambda'}(t,0)$ is invariant with respect to the shifts on $x^{n\varepsilon}$ $[x_i^{n\varepsilon}=0$ for all $i\in(\Lambda')^c]$. The above relation together with (5.7) implies that for any $\delta>0$,

$$\begin{aligned}\liminf_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^{\Lambda}(t,0)-vt|<\delta n\Big)\\ \geq \varepsilon\log\gamma+\liminf_{n\to+\infty}\frac{1}{n}\log\mathbb{P}\Big(\sup_{t\in[0,n\tau]}|X^{\Lambda'}(t,0)-vt|<\delta' n\Big)\end{aligned}$$

if $\varepsilon>\delta'>0$ are small enough, and hence (5.5) holds. Proposition 5.2 is proved. □

6. Fluid approximation. In this section we recall the results of [5], applied for localized Markov processes. We shall use these results for the proof of the lower large deviation bound in Section 9.

Given $\Lambda\subseteq\{1,\dots,N\}$, consider the Markov process $(X^{\Lambda}(t,y))$ on $\mathbb{Z}_+^{\Lambda,N}$, with initial state $X(0,y)=y$ and generator (5.1), and let

$$M(x)=(M_i(x))_{i=1}^N,\qquad x\in\mathbb{Z}_+^{\Lambda,N}$$

be a vector field of mean jumps of the process $(X^{\Lambda}(t,y))$,

$$M(x)=\sum_{y\in\mathbb{Z}_+^{\Lambda,N}}(y-x)q(y-x),\qquad x\in\mathbb{Z}_+^{\Lambda,N}.$$

Notice that the generator (5.1) is invariant with respect to the shifts on $z\in\mathbb{Z}_+^{\Lambda,N}$ if $z_i=0$ for all $i\in\Lambda^c$, and hence the projection of our Markov process onto

$$\mathbb{Z}_+^{\Lambda^c}=\{z\in\mathbb{Z}^N\colon z_\Lambda=0\text{ and }z_i\geq 0\text{ for all }i\in\Lambda^c\}$$

denoted by $(X_{\Lambda^c}^{\Lambda}(t,y))$, is a Markov process on $\mathbb{Z}_+^{\Lambda^c}$, with generator

$$\big(\mathscr{L}_{\Lambda^c}^{\Lambda}f\big)(x)=\sum_{y\in\mathbb{Z}_+^{\Lambda^c}}q_{\Lambda^c}(y-x)(f(y)-f(x)),\qquad x\in\mathbb{Z}_+^{\Lambda^c},$$

where

$$q_{\Lambda^c}(y)=\sum_{x\in\mathbb{Z}_+^{\Lambda,N}\colon x_{\Lambda^c}=y_{\Lambda^c}}q(x),\qquad y\in\mathbb{Z}_+^{\Lambda^c},$$

or using (1.1),

$$q_{\Lambda^c}(y)=\begin{cases}\lambda_i+\sum_{j\in\Lambda}\mu_j p_{ji}, & \text{if } y=\epsilon_i,\ i\in\Lambda^c,\\ \mu_i\Big(p_{i0}+\sum_{j\in\Lambda}p_{ij}\Big), & \text{if } y=-\epsilon_i,\ i\in\Lambda^c,\\ \mu_i p_{ij}, & \text{if } y=\epsilon_j-\epsilon_i,\ i,j\in\Lambda^c,\\ 0, & \text{otherwise.}\end{cases}$$

So the Markov process $(X^{\Lambda}_{\Lambda^c}(t, y))$ describes an open Jackson network with the set of nodes Λ^c, independent Poisson inputs with parameters

$$\lambda_i^{\Lambda} = \lambda_i + \sum_{j\in\Lambda} \mu_j p_{ji},$$

exponential service times with parameters μ_i and $P_{\Lambda^c,\Lambda^c} = (p_{ij})_{i,\,j\in\Lambda^c}$ is the corresponding transition matrix.

The Markov process $(X^{\Lambda}_{\Lambda^c}(t, y))$ is called an induced Markov process (induced Markov chain) corresponding to Λ (see [5]).

We say that the boundary $B_\Lambda = \{x \in \mathbb{R}^N : x_{\Lambda^c} = 0\}$ is attractive, if the induced Markov process $(X^{\Lambda}_{\Lambda^c}(t, y))$ is ergodic, and for the case where the above Markov process is ergodic, we denote by $\pi_{\Lambda^c}(x)$, $x \in \mathbb{Z}_+^{\Lambda^c}$ its stationary probabilities.

Remark that the ergodicity criteria for an open Jackson network [19] [see (1.3)] gives a necessary and sufficient condition for the ergodicity of $(X^{\Lambda}_{\Lambda^c}(t, y))$ and Jackson's product form (1.4) gives an explicit form for $\pi_{\Lambda^c}(x)$, $x \in \mathbb{Z}_+^{\Lambda^c}$.

Using the results of [22], Botvich and Zamyatin show that if the boundary B_Λ is attractive, then for any $x \in B_\Lambda$ almost surely,

$$\frac{1}{n} X^{\Lambda}(nt, [nx]) \to x + V^{\Lambda} t \quad \text{as } n \to +\infty,$$

uniformly on $t \in K$ for every compact set $K \subset \mathbb{R}_+$, where $V^{\Lambda} = (V_i^{\Lambda}; i = 1, \dots, N)$ is defined as follows

$$V_i^{\Lambda} = \begin{cases} \sum_{x\in\mathbb{Z}_+^{\Lambda^c}} \pi_{\Lambda^c}(x) M_i(x), & \text{if } i \in \Lambda, \\ 0, & \text{if } i \in \Lambda^c. \end{cases}$$

The vector V^{Λ} is called an induced vector relative to Λ.

The above result together with the ergodicity criteria for an open Jackson network [19] and the explicit form for the stationary probabilities $\pi_{\Lambda^c}(x)$, $x \in \mathbb{Z}_+^{\Lambda^c}$, imply the following statement (see [5]).

PROPOSITION 6.1. *Given* $\Lambda \subseteq \{1, \dots, N\}$, *consider*

$$\nu^{\Lambda}_{\Lambda^c} = (\lambda_{\Lambda^c} + \mu_\Lambda P_{\Lambda,\Lambda^c})(\mathbb{I}_{\Lambda^c,\Lambda^c} - P_{\Lambda^c,\Lambda^c})^{-1}$$

and $V^{\Lambda} = (V_i^{\Lambda})_{i=1}^N$ *with* $V^{\Lambda}_{\Lambda^c} = 0$ *and*

$$V^{\Lambda}_{\Lambda} = -\mu_\Lambda + \lambda_\Lambda + \mu_\Lambda P_{\Lambda,\Lambda} + \nu^{\Lambda}_{\Lambda^c} P_{\Lambda^c,\Lambda}. \tag{6.1}$$

Suppose that

$$\nu^{\Lambda}_{\Lambda^c} < \mu_{\Lambda^c}, \tag{6.2}$$

then for any $x \in B_\Lambda$ *almost surely*

$$\frac{1}{n} X^{\Lambda}(nt, [nx]) \to x + V^{\Lambda} t \quad \textit{as } n \to +\infty,$$

uniformly on compact sets.

The relation (6.2) is exactly the necessary and sufficient condition for the ergodicity of the induced Markov process $(X^{\Lambda}_{\Lambda^c}(t, y))$, and (6.1) gives an explicit form for the induced vector V^{Λ}.

We shall use this proposition for the proof of the lower large deviation bound.

7. Exponential change of measure. To prove lower and upper local large deviations bounds we shall use the method of exponential change of measure. This method consists in introducing a new probability measure by using a Radon–Nikodym factor in order to make what was originally "deviant" behavior look like typical behavior.

In this section we recall the definition of exponential change of measure and we show that a new random process relative to a new probability measure as well as the initial process describes the queue length process of an open Jackson network (Lemma 7.2). Using this together with the fluid approximation for the above processes given by Proposition 6.1, we describe the typical behavior of a new process in Proposition 7.1.

Consider a Markov process $(X^{\Lambda}(t, y))$, with the set of states $\mathbb{Z}_+^{\Lambda, N}$, initial state $X(0, y) = y$ and generator (5.1). Let $\mathbb{P}_{y,\Lambda}$ be the distribution of the above random process; denote by $\mathbb{E}_{y,\Lambda}$ the expectation with respect to $\mathbb{P}_{y,\Lambda}$ and let $\{\mathscr{F}_t^{y,\Lambda}\}_{t\in\mathbb{R}_+}$ be the natural filtration,

$$\mathscr{F}_t^{y,\Lambda} = \sigma(X^{\Lambda}(s, y), s \le t), \qquad t \in \mathbb{R}_+.$$

Given $\alpha \in \mathbb{R}^N$ and $y \in \mathbb{Z}_+^{\Lambda, N}$, define

$$\mathscr{M}_y(\alpha, t) = \exp\left\{\langle \alpha, X^{\Lambda}(t, y) - y\rangle - \int_0^t R_{\Lambda}(\alpha, X^{\Lambda}(s, y))\,ds\right\}, \qquad t \in \mathbb{R}_+,$$

where

$$R_{\Lambda}(\alpha, x) = \sum_{z\in\mathbb{Z}_+^{\Lambda, N}} q(z-x)(\exp(\langle \alpha, z-x\rangle) - 1), \qquad x \in \mathbb{Z}_+^{\Lambda, N}. \tag{7.1}$$

LEMMA 7.1. *For any $\alpha \in \mathbb{R}^N$ and $y \in \mathbb{Z}_+^{\Lambda, N}$, $\mathscr{M}_y(\alpha, t)$ is a martingale relative to $(\{\mathscr{F}_t^{y,\Lambda}\}_{t\in\mathbb{R}_+}, \mathbb{P}_{y,\Lambda})$ with*

$$\mathbb{E}_{y,\Lambda}(\mathscr{M}_y(\alpha, t)) \equiv 1. \tag{7.2}$$

PROOF. Indeed, given $\alpha \in \mathbb{R}^N$, for any $y \in \mathbb{Z}_+^{\Lambda, N}$,

$$\frac{d}{dt}\mathbb{E}_{y,\Lambda}(\mathscr{M}_y(\alpha, t))\bigg|_{t=0} = \sum_{z\in\mathbb{Z}_+^{\Lambda, N}} q(z-y)(\exp(\langle \alpha, z-y\rangle) - 1) - R_{\Lambda}(\alpha, y) = 0,$$

and hence, using Markov property of the process $(X^{\Lambda}(t, y))_{t\in\mathbb{R}_+}$ we get

$$\frac{d}{dt}\mathbb{E}_{y,\Lambda}(\mathscr{M}_y(\alpha, t)) = 0 \quad \text{for all } t \ge 0.$$

The above implies that

$$\mathbb{E}_{y,\Lambda}(\mathscr{M}_y(\alpha,\ t)) \equiv \mathbb{E}_{y,\Lambda}(\mathscr{M}_y(\alpha,\ 0)) = 1$$

and therefore (7.2) holds.

From (7.2) using again Markov property of the process $(X^\Lambda(t,\ y))$ we get

$$\mathbb{E}_{y,\Lambda}\big(\mathscr{M}_y(\alpha,t)|\mathscr{F}_s^{y,\Lambda}\big) = \mathscr{M}_y(\alpha,s) \quad \text{for all } t \ge s \ge 0,$$

and therefore, $\mathscr{M}_y(\alpha,t)$ is a martingale relative to $\big(\{\mathscr{F}_t^{y,\Lambda}\}_{t\in\mathbb{R}_+}, \mathbb{P}_{y,\Lambda}\big)$. Lemma 7.1 is proved. □

Using Lemma 7.1, we define new probability measures $\mathbb{P}_{y,\Lambda}^{(\alpha)}$ on

$$\mathscr{F}^{y,\Lambda} = \bigcup_{t\ge 0} \mathscr{F}_t^{y,\Lambda},$$

such that for all $t \ge 0$ and $A \in \mathscr{F}_t^{y,\Lambda}$,

$$\mathbb{P}_{y,\Lambda}^{(\alpha)}(A) = \mathbb{E}_{y,\Lambda}(\mathbb{1}_A \mathscr{M}_y(\alpha,t)) \tag{7.3}$$

where $\mathbb{1}_A$ denotes the indicator of A.

Denote by $\mathbb{E}_{y,\Lambda}^{(\alpha)}$ the expectation with respect of the measure $\mathbb{P}_{y,\Lambda}^{(\alpha)}$. Then (7.3) implies that for all $t \ge 0$ and $A \in \mathscr{F}_t^{y,\Lambda}$,

$$\mathbb{P}_{y,\Lambda}(A) = \mathbb{E}_{y,\Lambda}^{(\alpha)}(\mathbb{1}_A(\mathscr{M}_y(\alpha,t))^{-1}). \tag{7.4}$$

As usual $\mathbb{P}_{y,\Lambda}^{(\alpha)}$ is called an exponential change of the measure $\mathbb{P}_{y,\Lambda}$ or twisted distribution. The Markov process $(X^\Lambda(t,\ y))$ relative to $\mathbb{P}_{y,\Lambda}^{(\alpha)}$ is called twisted Markov process.

The following lemma shows that the above change of measure corresponds to simply changing the jump rates from $q(y)$ to $e^{\langle\alpha,\,y\rangle}q(y)$.

LEMMA 7.2. *For any $\alpha \in \mathbb{R}^N$, the twisted Markov process $(X^\Lambda(t,\ y))$ relative to $\mathbb{P}_{y,\Lambda}^{(\alpha)}$, describes an open Jackson network with N nodes, where the nodes $j \in \Lambda$ are free, independent Poisson inputs with parameters*

$$\lambda_i(\alpha) = \lambda_i e^{\alpha_i}, \qquad i \in \{1,\dots,N\}, \tag{7.5}$$

exponential service times with parameters

$$\mu_i(\alpha) = \mu_i e^{-\alpha_i}\left(\sum_{j=1}^N p_{ij}e^{\alpha_j} + p_{i0}\right), \qquad i \in \{1,\dots,N\} \tag{7.6}$$

and routing matrix $(p_{ij}(\alpha)\colon i, j = 0, \dots, N)$ *where*

$$
\begin{aligned}
&p_{ij}(\alpha) = \frac{p_{ij}e^{\alpha_j}}{\sum_{k=1}^N p_{ik}e^{\alpha_k} + p_{i0}}, \qquad i, j \in \{1, \dots, N\},\\
(7.7)\qquad &p_{i0}(\alpha) = \frac{p_{i0}}{\sum_{k=1}^N p_{ik}e^{\alpha_k} + p_{i0}}, \qquad i \in \{1, \dots, N\},\\
&p_{0i}(\alpha) = p_{0i} = \begin{cases} 1, & \textit{if } i = 0,\\ 0, & \textit{otherwise}. \end{cases}
\end{aligned}
$$

PROOF. Indeed, the Markov property of the twisted process is a simple consequence of the Markov property of the original process $(X^\Lambda(t, y))$ (see, e.g., the section on Doob's h-transform in [27] or [25]). Moreover, relation (7.3) implies that for any compactly supported function $f\colon \mathbb{Z}_+^{\Lambda, N} \to \mathbb{R}$,

$$
\begin{aligned}
&\mathbb{E}_{y,\Lambda}^{(\alpha)}(f(X^\Lambda(t, y)))\\
&\quad = \mathbb{E}_{y,\Lambda}\Big(f\big(X^\Lambda(t, y)\big)\exp\Big\{\langle \alpha, X^\Lambda(t, y) - y\rangle - \int_0^t R_\Lambda(\alpha, X^\Lambda(s, y)), ds\Big\}\Big),
\end{aligned}
$$

and hence

$$
\begin{aligned}
&\frac{d}{dt}\mathbb{E}_{y,\Lambda}^{(\alpha)}\big(f(X^\Lambda(t, y))\big)\Big|_{t=0}\\
&\qquad = \sum_{x \in \mathbb{Z}_+^{\Lambda, N}} q(x - y)\big(f(x)\exp(\langle \alpha, x - y\rangle) - f(y)\big) - f(y)R_\Lambda(\alpha, y).
\end{aligned}
$$

Using now (7.1) we get

$$
\frac{d}{dt}\mathbb{E}_{y,\Lambda}^{(\alpha)}\big(f(X^\Lambda(t, y))\big)\Big|_{t=0} = \sum_{x \in \mathbb{Z}_+^{\Lambda, N}} q(x - y)\exp(\langle \alpha, x - y\rangle)(f(x) - f(y))
$$

and therefore, $\mathbb{P}_{y,\Lambda}^{(\alpha)}$ is a distribution of a Markov process on $\mathbb{Z}_+^{\Lambda, N}$ with generator

$$
(\mathscr{L}^{\Lambda,\alpha} f)(x) = \sum_{z \in \mathbb{Z}_+^{\Lambda, N}} q(z - x)\exp(\langle \alpha, z - x\rangle)(f(z) - f(x)), \qquad x \in \mathbb{Z}_+^{\Lambda, N}.
$$

Moreover, for $i, j \in \{1, \dots, N\}$,

$$
(7.8)\qquad q(y)e^{\langle \alpha, v\rangle} = \begin{cases} \lambda_i(\alpha), & \text{if } y = \epsilon^i,\\ \mu_i(\alpha)p_{i0}(\alpha), & \text{if } y = -\epsilon^i,\\ \mu_i(\alpha)p_{ij}(\alpha), & \text{if } y = \epsilon^j - \epsilon^i,\\ 0, & \text{otherwise}, \end{cases}
$$

and the matrix $(p_{ij}(\alpha);\ i, j = 0, \dots, N)$ is obviously stochastic. Finally, comparison of (7.8) with (1.1) completes the proof of our lemma. □

We are ready now to describe the typical behavior of a twisted Markov process relative to $\mathbb{P}^{(\alpha)}_{y,\Lambda}$.

Given $\alpha \in \mathbb{R}^N$, consider the vectors

$$\nu^{\Lambda}_{\Lambda^c}(\alpha) = (\lambda_{\Lambda^c}(\alpha) + \mu_{\Lambda}(\alpha) P_{\Lambda\Lambda^c}(\alpha))(\mathbb{I} - P_{\Lambda^c\Lambda^c}(\alpha))^{-1}, \tag{7.9}$$

and $V^{\Lambda}(\alpha) = (V^{\Lambda}_i(\alpha);\ i = 1, \dots, N)$ with $V^{\Lambda}_i(\alpha) = 0$, for $i \in \Lambda^c$ and

$$V^{\Lambda}_{\Lambda}(\alpha) = -\mu_{\Lambda}(\alpha) + \lambda_{\Lambda}(\alpha) + \mu_{\Lambda}(\alpha) P_{\Lambda\Lambda}(\alpha) + \nu^{\Lambda}_{\Lambda^c}(\alpha) P_{\Lambda^c\Lambda}(\alpha), \tag{7.10}$$

and notice that under Assumption A, the matrix $(p_{ij}(\alpha);\ i, j = 1, \dots, N)$ as well as the matrix $(p_{ij};\ i, j = 1, \dots, N)$ has a spectral radius less than unity. Then Lemma 7.2 together with Proposition 6.1 imply the following statement.

PROPOSITION 7.1. *Suppose that*

$$\nu^{\Lambda}_{\Lambda^c}(\alpha) < \mu_{\Lambda^c}(\alpha).$$

Then

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}^{(\alpha)}_{0,\Lambda}\left(\sup_{t\in[0,n]} |X^{\Lambda}(0,t) - V^{\Lambda}(\alpha)t| < \delta n\right) = 0$$

for all $\delta > 0$.

8. The properties of the functions $l^{\Lambda}(\cdot)$. In this section we study the properties of the functions $l^{\Lambda}(\cdot)$. In particular, a relationship between the above functions and the fluid approximation (typical behavior) of twisted Markov processes is established. These results will be used for the proof of the local lower large deviation bound.

Recall that

$$l^{\Lambda}(v) = \sup_{\alpha\in\mathscr{B}_{\Lambda}} \{\langle\alpha, v\rangle - R(\alpha)\}, \Lambda \subseteq \{1, \dots, N\},$$

where

$$R(\alpha) = \sum_{y\in\mathbb{Z}^N} q(y)(e^{\langle\alpha, y\rangle} - 1)$$

and

$$\mathscr{B}_{\Lambda} = \left\{\alpha \in \mathbb{R}^N : \alpha_j \le \log\left(\sum_{i=1}^N p_{ji} e^{\alpha_i} + p_{j0}\right) \text{ for all } j \in \Lambda^c\right\}.$$

We start by rewriting the values $l^{\Lambda}(v)$ for given $\Lambda \subseteq \{1, \dots, N\}$ and $v \in \mathbb{R}^N$, $v_{\Lambda^c} = 0$, as a maximum of a strictly concave function $\beta \to \langle\beta, v\rangle - R(\alpha(\beta))$ in the convex set

$$\mathbb{R}^{\Lambda,N}_{\le 0} = \{\beta \in \mathbb{R}^N : \beta_i \le 0 \text{ for all } i \in \Lambda^c\}.$$

The following proposition gives a diffeomorphism $\alpha(\cdot)$ from $\mathbb{R}^{\Lambda,N}_{\le 0}$ to $\mathscr{B}_{\Lambda}$ for which the function $R(\alpha(\cdot))$ is strictly convex.

PROPOSITION 8.1. *Let* $\Lambda \subseteq \{1, \ldots, N\}$. *Consider the system*

$$\begin{cases} \alpha_i = \beta_i, & \text{for } i \in \Lambda, \\ \alpha_j - \log\left(\sum_{k=1}^{N} p_{jk} e^{\alpha_k} + p_{j0}\right) = \beta_j, & \text{for } j \in \Lambda^c, \end{cases} \tag{8.1}$$

and let

$$\mathbb{R}^{\Lambda, N}_{<\varepsilon} = \{\beta \in \mathbb{R}^N : \beta_j < \varepsilon \text{ for all } j \in \Lambda^c\}.$$

Then there exists $\varepsilon_0 > 0$, *such that*:

(i) *For any* $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$, *the system* (8.1) *has a unique solution* $\alpha(\beta)$,

$$\begin{cases} \alpha_i(\beta) = \beta_i, & \text{for } i \in \Lambda, \\ \alpha_i(\beta) = \log\left\{\sum_{j \in \Lambda} m^{\Lambda}_{ij}(\beta) e^{\beta_j} + m^{\Lambda}_{i0}(\beta)\right\}, & \text{for } i \in \Lambda^c, \end{cases} \tag{8.2}$$

where given $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ *and* $i, j \in \{0, 1, \ldots, N\}$, $i \neq 0$, *we denote*

$$m^{\Lambda}_{ij}(\beta) = p_{ij} e^{\beta_i} + \sum_{n \geq 1} \sum_{j_1, \ldots, j_n \in \Lambda^c} p_{ij_1} p_{j_1 j_2} \cdots p_{j_n j} \exp\left(\beta_i + \sum_{k=1}^{n} \beta_{j_k}\right).$$

(ii) $\alpha(\cdot), R(\alpha(\cdot)) \in C^\infty(\mathbb{R}^{\Lambda, N}_{<\varepsilon_0})$ *and the function* $R(\alpha(\cdot))$ *is strictly convex everywhere on* $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$.

(iii) *For any* $v \in \mathbb{R}^N$ *such that* $v_{\Lambda^c} = 0$,

$$l^{\Lambda}(v) = \sup_{\beta \in \mathbb{R}^{\Lambda, N}_{\leq 0}} \{\langle \beta, v \rangle - R(\alpha(\beta))\}. \tag{8.3}$$

(iv) *For every* $c \in \mathbb{R}_+$ *and* $v \in \mathbb{R}^N$ *such that* $v_{\Lambda^c} = 0$, *the set*

$$\{\beta \in \mathbb{R}^{\Lambda, N}_{\leq 0} : R(\alpha(\beta)) - \langle \beta, v \rangle \leq c\}$$

is a compact subset of $\mathbb{R}^N$.

We shall prove this proposition in the Appendix.

The following proposition relates the function $R(\alpha(\cdot))$ and the fluid approximation of the twisted Markov process $(X^{\Lambda}(t, y))$ relative to $\mathbb{P}^{(\alpha)}_{y, \Lambda}$. Recall that this fluid approximation can be described by using the vectors $\nu^{\Lambda}_{\Lambda^c}(\alpha)$ and $V^{\Lambda}(\alpha)$ which are defined by (7.9) and (7.10), respectively.

PROPOSITION 8.2. *Given* $\Lambda \subseteq \{1, \ldots, N\}$, *let* $\alpha(\beta)$ *be the unique solution of the system* (8.1). *Consider*

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta)) = \left(\frac{\partial}{\partial \beta_i} R(\alpha(\beta))\right)_{i \in \Lambda^c}$$

and

$$\nabla_{\beta_\Lambda} R(\alpha(\beta)) = \left(\frac{\partial}{\partial \beta_i} R(\alpha(\beta))\right)_{i\in\Lambda}.$$

Then for any $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$,

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta)) = \nu^{\Lambda}_{\Lambda^c}(\alpha(\beta)) - \mu_{\Lambda^c}(\alpha(\beta)) \tag{8.4}$$

and

$$\nabla_{\beta_\Lambda} R(\alpha(\beta)) = V^{\Lambda}_{\Lambda}(\alpha(\beta)). \tag{8.5}$$

PROOF. Indeed, using (2.4) together with (8.1) we get

$$\begin{aligned} R(\alpha(\beta)) = {} & \sum_{j\in\Lambda^c} \mu_j(\exp(-\beta_j) - 1) + \sum_{j\in\Lambda^c} \lambda_j(\exp(\alpha_j(\beta)) - 1) \\ & + \sum_{j\in\Lambda} \lambda_j(\exp(\beta_j) - 1) \\ & + \sum_{j\in\Lambda} \mu_j\bigg(\sum_{k\in\Lambda} p_{jk}\exp(\beta_k - \beta_j) \\ & \qquad + p_{j0}\exp(-\beta_j) + \sum_{k\in\Lambda^c} p_{jk}\exp(\alpha_k(\beta) - \beta_j) - 1\bigg). \end{aligned} \tag{8.6}$$

for any $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$. Consider

$$a_{ij}(\beta) = \frac{\partial}{\partial \beta_j}\alpha_i(\beta), \qquad i, j \in \{1, \dots, N\},$$

let $A_{\Lambda^c\Lambda^c}(\beta) = (a_{ij}(\beta);\ i, j \in \Lambda^c)$, and $A_{\Lambda^c\Lambda}(\beta) = (a_{ij}(\beta);\ i \in \Lambda^c,\ j \in \Lambda)$. Then by (8.1),

$$A_{\Lambda^c\Lambda^c}(\beta) = (\mathbb{I}_{\Lambda^c\Lambda^c} - P_{\Lambda^c\Lambda^c}(\alpha(\beta)))^{-1}$$

and

$$A_{\Lambda^c\Lambda}(\beta) = (\mathbb{I}_{\Lambda^c\Lambda^c} - P_{\Lambda^c\Lambda^c}(\alpha(\beta)))^{-1} P_{\Lambda^c\Lambda}(\alpha(\beta)).$$

Using the above expression for $A_{\Lambda^c\Lambda^c}(\beta)$ and $A_{\Lambda^c\Lambda}(\beta)$ together with (8.6) we get (8.4) and (8.5), and therefore, Proposition 8.2 holds. □

Proposition 8.1 and Proposition 8.2 imply the following statement.

PROPOSITION 8.3. *Let* $\Lambda \subseteq \{1, \dots, N\}$ *and* $v \in \mathbb{R}^N$ *such that* $v_{\Lambda^c} = 0$.

(i) *Then there exists* $\alpha^v \in \mathscr{B}_\Lambda$ *such that*

$$l^{\Lambda}(v) = \langle \alpha^v, v\rangle - R(\alpha^v); \tag{8.7}$$

α^v *is a unique point of a local maximum of the function* $\alpha \to \langle \alpha, v\rangle - R(\alpha)$ *in* $\mathscr{B}_\Lambda$ *(and hence* α^v *is unique for given* Λ *and* v*).*

(ii) $V^{\Lambda}(\alpha^v) = v$ *and* $\nu^{\Lambda}_{\Lambda^c}(\alpha^v) \le \mu_{\Lambda^c}(\alpha^v)$.

(iii) *Suppose, moreover, that for some* $i \in \Lambda^c$, *either* $\nu_i^{\Lambda}(\alpha^v) - \mu_i(\alpha^v) = 0$ *or* $\alpha_i^v < \log(\sum_{j=1}^N p_{ij} e^{\alpha_j^v} + p_{i0})$, *then*

$$l^{\Lambda}(v) = l^{\Lambda \cup \{i\}}(v). \tag{8.8}$$

PROOF. Let $\alpha(\beta)$ be the unique solution of the system (8.1) for given Λ. Proposition 8.1 shows that $\alpha(\cdot)$ is a homeomorphism from $\mathbb{R}_{\le 0}^{\Lambda, N}$ to $\mathscr{B}_{\Lambda}$ and for $\alpha = \alpha(\beta)$,

$$\langle \alpha, v \rangle = \langle \beta, v \rangle$$

because $v_{\Lambda^c} = 0$ and $\alpha_{\Lambda} = \beta_{\Lambda}$. The above implies that

$$l^{\Lambda}(v) = \sup_{\beta \in \mathbb{R}_{\le 0}^{\Lambda, N}} \{\langle \beta, v \rangle - R(\alpha(\beta))\}.$$

Proposition 8.1 proves moreover that the function $R(\alpha(\cdot))$ is strictly convex on $\mathbb{R}_{\le 0}^{\Lambda, N}$ and set $\{\beta \in \mathbb{R}_{\le 0}^{\Lambda, N} \colon R(\alpha(\beta)) - \langle \beta, v \rangle \le 0\}$ is compact and nonempty, because $R(\alpha(0)) = 0$. Hence, there exists a unique $\beta^v \in \mathbb{R}_{\le 0}^{\Lambda, N}$ such that

$$l^{\Lambda}(v) = \langle \beta^v, v \rangle - R(\alpha(\beta^v)). \tag{8.9}$$

Clearly, β^v is a unique point of a local maximum of the function $\beta \to \langle \beta, v \rangle - R(\alpha(\beta))$ in $\mathbb{R}_{\le 0}^{\Lambda, N}$.

Consider now $\alpha^v = \alpha(\beta^v) \in \mathscr{B}_{\Lambda}$. Since $\alpha(\cdot)$ is a homeomorphism from $\mathbb{R}_{\le 0}^{\Lambda, N}$ to $\mathscr{B}_{\Lambda}$ then α^v achieves the maximum of the function $\langle \alpha, v \rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda}$,

$$l^{\Lambda}(v) = \langle \alpha^v, v \rangle - R(\alpha^v)$$

and, moreover, α^v is a unique point of a local maximum of this function in $\mathscr{B}_{\Lambda}$. The first part of our proposition is therefore verified with $\alpha^v = \alpha(\beta^v)$.

To prove the second part of Proposition 8.3, let us notice that

$$\nabla_{\beta_{\Lambda}} R(\alpha(\beta))|_{\beta = \beta^v} = v_{\Lambda} \tag{8.10}$$

and

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta))|_{\beta = \beta^v} \le 0 \tag{8.11}$$

because β^v is a point of a local maximum of the function $\beta \to \langle \beta, v \rangle - R(\alpha(\beta))$ in

$$\mathbb{R}_{\le 0}^{\Lambda, N} = \{\beta \in \mathbb{R}^N \colon \beta_i \le 0 \text{ for all } i \in \Lambda^c\}.$$

But Proposition 8.1 shows that

$$\nabla_{\beta_{\Lambda}} R(\alpha(\beta))|_{\beta = \beta^v} = V_{\Lambda}^{\Lambda}(\alpha(\beta^v)) = V_{\Lambda}^{\Lambda}(\alpha^v)$$

and

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta))|_{\beta = \beta^v} = \nu_{\Lambda^c}^{\Lambda}(\alpha(\beta^v)) - \mu_{\Lambda^c}(\alpha(\beta^v)) = \nu_{\Lambda^c}^{\Lambda}(\alpha^v) - \mu_{\Lambda^c}(\alpha^v)$$

and hence, relations (8.10) and (8.11) imply that $V_{\Lambda}^{\Lambda}(\alpha^v) = v_{\Lambda}$ and $\nu_{\Lambda^c}^{\Lambda}(\alpha^v) \le \mu_{\Lambda^c}(\alpha^v)$. The second part of Proposition 8.3 is therefore proved.

Suppose now that $\alpha_i^v < \log(\sum_{j=1}^N p_{ij}\exp(\alpha_j^v) + p_{i0})$ for some $i \in \Lambda^c$, then α^v is a point of a local maximum of the function $\alpha \to \langle \alpha, v\rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda\cup\{i\}}$, and since it is unique, then

$$l^{\Lambda\cup\{i\}}(v) = \sup_{\alpha\in\mathscr{B}_{\Lambda\cup\{i\}}} \{\langle \alpha, v\rangle - R(\alpha)\} = \langle \alpha^v, v\rangle - R(\alpha^v) = l^\Lambda(v)$$

and the relation (8.8) holds.

Finally, if for some $i \in \Lambda^c$,

$$\nu_i^\Lambda(\alpha^v) - \mu_i(\alpha^v) = 0,$$

then because of Proposition 8.2,

$$\frac{\partial}{\partial \beta_i} R(\alpha(\beta))\Big|_{\beta=\beta^v} = 0,$$

and since the function $R(\alpha(\cdot))$ is convex everywhere on $\mathbb{R}_{\varepsilon_0}^{\Lambda, N}$, the above relation together with (8.10) and (8.11) imply that β^v is a point of a local maximum of the function $\beta \to \langle \beta, v\rangle - R(\alpha(\beta))$ in

$$D = \{\beta \in \mathbb{R}^\Lambda \times \mathbb{R}^{\Lambda^c} \colon \beta_i < \varepsilon_0 \text{ and } \beta_j \le 0 \text{ for } j \in \Lambda^c, j \ne i\}.$$

Observe that the mapping $\beta \to \alpha(\beta)$ is a homeomorphism from D to

$$\mathscr{B} = \Bigg\{\alpha \in \mathbb{R}^\Lambda \times \mathbb{R}^{\Lambda^c} \colon \alpha_i < \varepsilon_0 + \log\left(\sum_{k=1}^N p_{ik}e^{\alpha_k} + p_{i0}\right) \text{ and}$$
$$\alpha_j \le \log\left(\sum_{k=1}^N p_{jk}e^{\alpha_k} + p_{j0}\right) \text{ for } j \in \Lambda^c, j \ne i \Bigg\}$$

and the set $\mathscr{B}$ is open in $\mathscr{B}_{\Lambda\cup\{j\}}$ with respect to the topology induced by Euclidean topology in $\mathbb{R}^N$. This proves that $\alpha^v = \alpha(\beta^v)$ is the point of a local maximum of the function $\alpha \to \langle \alpha, v\rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda\cup\{j\}}$, and since this point is unique, the relation (8.8) holds. Proposition 8.3 is proved. □

9. Local large deviation bounds. Given $\Lambda \subseteq \{1, \dots, N\}$, consider a Markov process $(X^\Lambda(t, y))$ with the set of states $\mathbb{Z}_+^{\Lambda, N}$, initial state $X^\Lambda(0, y) = y \in \mathbb{Z}_+^{\Lambda, N}$ and generator (5.1). The main result of this section is the following proposition. It completes the proof of Theorem 1.

PROPOSITION 9.1. *For any $\tau > 0$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, the following relation holds*:

$$l^\Lambda(v) = -\frac{1}{\tau}\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{1}{n}\log \mathbb{P}_{0,\Lambda}\left(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0) - vt| < \delta n\right)$$
$$= -\frac{1}{\tau}\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log \mathbb{P}_{0,\Lambda}\left(\sup_{t\in[0,n\tau]}|X^\Lambda(t,0) - vt| < \delta n\right).$$

PROOF. Let $v \in \mathbb{R}^N$ and $v_{\Lambda^c} = 0$. Denote

$$A_{n\delta} = \left\{ \sup_{t\in[0,\,n\tau]} |X^\Lambda(t,0) - vt| < \delta n \right\}.$$

To prove our theorem we have to verify the upper large deviation bound,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{0,\,\Lambda}(A_{n\delta}) \le -\tau l^\Lambda(v) \tag{9.1}$$

and the lower large deviation bound,

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{0,\,\Lambda}(A_{n\delta}) \ge -\tau l^\Lambda(v). \tag{9.2}$$

We start with the proof of the upper large deviation bound. Let $\mathbb{P}^{(\alpha)}_{0,\,\Lambda}$ be the exponential change of the measure $\mathbb{P}_{0,\,\Lambda}$ for $\alpha \in \mathbb{R}^N$ (see Section 7). Then using (7.4) we get

$$\begin{aligned} &\mathbb{P}_{0,\,\Lambda}(A_{n\delta}) \\ &\quad = \mathbb{E}^{(\alpha)}_{0,\,\Lambda}\left(\mathbb{1}_{\{A_{n\delta}\}} \exp\left\{ -\langle \alpha, X^\Lambda(n\tau,0)\rangle + \int_0^{n\tau} R_\Lambda(\alpha, X^\Lambda(t,0))\,ds \right\}\right), \end{aligned} \tag{9.3}$$

where

$$R_\Lambda(\alpha,\, z) = \sum_{y\in\mathbb{Z}_+^{\Lambda,\,N}} q(y-z)(\exp(\langle \alpha, y-z\rangle) - 1), \qquad z \in \mathbb{Z}_+^{\Lambda,\,N}.$$

Furthermore, using (1.1) it follows that

$$\begin{aligned} &R_\Lambda(\alpha, z) \\ &\quad = R(\alpha) - \sum_{i\in\Lambda^c} \mathbb{1}_{\{z_i=0\}} \mu_i \left(\sum_{j=1}^N p_{ij} \exp(\alpha_j - \alpha_i) + p_{i0}\exp(-\alpha_i) - 1 \right), \end{aligned} \tag{9.4}$$

and hence, for any α such that

$$\alpha_i \le \log\left(\sum_{j=1}^N p_{ij} e^{\alpha_j} + p_{i0} \right) \quad \text{for all } i \in \Lambda^c \quad (\Leftrightarrow \alpha \in \mathscr{B}_\Lambda),$$

we have

$$R_\Lambda(\alpha, z) \le R(\alpha) \quad \text{for all } z \in \mathbb{Z}_+^{\Lambda,\,N}.$$

The above inequality, together with (9.3), gives

$$\mathbb{P}_{0,\,\Lambda}(A_{n\delta}) \le \mathbb{E}^{(\alpha)}_{0,\,\Lambda}\big(\mathbb{1}_{\{A_{n\delta}\}} \exp\{ -\langle \alpha, X^\Lambda(n\tau,0)\rangle + n\tau R(\alpha)\}\big), \tag{9.5}$$

for all $\alpha \in \mathscr{B}_\Lambda$. Notice now that for any trajectory $X^\Lambda(t,0); t \in \mathbb{R}_+$, for which $A_{n\delta}$ holds, we have

$$|X^\Lambda(n\tau,0) - vn\tau| < \delta n, \tag{9.6}$$

and hence (9.5) implies that for $\alpha \in \mathscr{B}_\Lambda$,

$$\log \mathbb{P}_{0,\Lambda}(A_{n\delta}) \le \delta|\alpha|n - n\tau(\langle \alpha, v\rangle - R(\alpha)).$$

The above relation gives the upper large deviation bound (9.1).

Let us prove the lower large deviation bound (9.2). For this we shall use an induction with respect to $k_\Lambda = |\Lambda^c| = N - |\Lambda|$. If $k_\Lambda = 0$, that is $\Lambda = \{1, \dots, N\}$, then our Markov process $(X^{\{1,\dots,N\}}(t, 0))$ is a homogeneous random walk in $\mathbb{Z}^N$ with the generator

$$\mathscr{L} f(x) = \sum_{z \in \mathbb{Z}^N} q(y)(f(y) - f(x)),$$

and the inequality (9.2) for that follows from the large deviation principle for homogeneous random walks in $\mathbb{Z}^N$.

Consider now $\Lambda \subseteq \{1, \dots, N\}$, such that $\Lambda^c \neq \varnothing$ and suppose that for all $j \in \Lambda^c$,

$$\begin{aligned}\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{0,\Lambda\cup\{j\}}\Big(\sup_{t\in[0,n\tau]} |X^{\Lambda\cup\{j\}}(0,t) - vt| < \delta n\Big) \\ \ge -\tau l^{\Lambda\cup\{j\}}(v).\end{aligned} \tag{9.7}$$

For $\Lambda \subset \Lambda' \subseteq \{1, \dots, N\}$, Proposition 5.2 yields

$$\begin{aligned}&\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{0,\Lambda}\Big(\sup_{t\in[0,n\tau]} |X^{\Lambda}(0,t) - vt| < \delta n\Big) \\ &\quad\ge \lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{0,\Lambda'}\Big(\sup_{t\in[0,n\tau]} |X^{\Lambda'}(0,t) - vt| < \delta n\Big)\end{aligned}$$

and hence, if

$$l^\Lambda(v) = l^{\Lambda\cup\{j\}}(v) \quad \text{for some } j \in \Lambda^c,$$

then our lower large deviation bound (9.2) follows from (9.7). Otherwise,

$$l^\Lambda(v) < l^{\Lambda\cup\{j\}}(v) \quad \text{for all } j \in \Lambda^c,$$

and because of Proposition 8.3, there exists a unique $\alpha^v \in \mathscr{B}_\Lambda$ such that

$$l^\Lambda(v) = \langle \alpha^v, v\rangle - R(\alpha^v), \tag{9.8}$$

$$\alpha_j^v = \log\Big(\sum_{k=1}^N p_{jk} e^{\alpha_k^v} + p_{j0}\Big) \quad \text{for all } j \in \Lambda^c, \tag{9.9}$$

$$\nu_j^\Lambda(\alpha^v) - \mu_j(\alpha^v) < 0 \quad \text{for all } j \in \Lambda^c \tag{9.10}$$

and

$$V^\Lambda(\alpha^v) = v. \tag{9.11}$$

Relation (9.9) together with (9.4) implies that

$$R_\Lambda(\alpha^v, z) = R(\alpha^v) \quad \text{for all } z \in \mathbb{Z}_+^{\Lambda, N},$$

and hence, using (9.3), we get

$$\mathbb{P}_{0,\Lambda}(A_{n\delta}) = \mathbb{E}_{0,\Lambda}^{(\alpha^v)}\big(\mathbb{1}_{\{A_{n\delta}\}} \exp\{-\langle \alpha^v, X^{\{\Lambda\}}(n\tau, 0)\rangle + n\tau R(\alpha^v)\}\big).$$

Furthermore, using (9.6), it follows that

$$\log \mathbb{P}_{0,\Lambda'}(A_{n\delta}) \geq n\tau(R(\alpha^v) - \langle \alpha^v, v\rangle) - \delta|\alpha^v|n + \log \mathbb{P}_{0,\Lambda}^{(\alpha^v)}(A_{n\delta}). \tag{9.12}$$

Finally, Proposition 7.1 together with (9.10) and (9.11) yields

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}_{0,\Lambda'}^{(\alpha^v)}(A_{n\delta}) = 0 \quad \text{for all } \delta > 0,$$

and hence, using (9.12) we obtain

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}_{0,\Lambda'}(A_{n\delta}) \geq \tau(R(\alpha^v) - \langle \alpha^v, v\rangle) - \delta|\alpha^v|.$$

The above relation together with (9.8) implies the local lower large deviation bound. Proposition 9.1 is therefore proved. □

10. The explicit expression of the local rate function. In this section we prove Theorem 2. Given $\Lambda \subseteq \{1, \dots, N\}$, consider the set

$$\mathscr{D}_\Lambda = \left\{\alpha \in \mathbb{R}^N: \alpha_j = \log\left(\sum_{k=1}^N p_{jk} e^{\alpha_k} + p_{j0}\right) \text{ for all } j \in \Lambda^c\right\}.$$

It is clear that $\alpha \in \mathscr{D}_\Lambda$ if and only if

$$\alpha = \alpha(\beta)\Big|_{\beta_\Lambda = \alpha_\Lambda,\, \beta_{\Lambda^c} = 0},$$

where $\alpha(\cdot)$ is the unique solution of the system (8.1) for given Λ and hence, for $\alpha \in \mathscr{D}_\Lambda$, one can rewrite the function $R(\alpha)$ and the vector $\nu_{\Lambda^c}^\Lambda(\alpha)$ in terms of α_Λ as follows:

$$R(\alpha)|_{\alpha\in D_\Lambda} = R(\alpha(\beta))\Big|_{\beta_\Lambda = \alpha_\Lambda,\, \beta_{\Lambda^c} = 0}, \tag{10.1}$$

$$\nu_{\Lambda^c}^\Lambda(\alpha) = \nu_{\Lambda^c}^\Lambda(\alpha(\beta))\Big|_{\beta_\Lambda = \alpha_\Lambda,\, \beta_{\Lambda^c} = 0}. \tag{10.2}$$

The following lemma gives the explicit form for them.

LEMMA 10.1. *For any $\alpha \in \mathscr{D}_\Lambda$, $R(\alpha) = H_\Lambda(\alpha_\Lambda)$ where*

$$\begin{aligned} H_\Lambda(\alpha_\Lambda) = {} & \sum_{i\in\Lambda}\left(\nu_i - \sum_{j\in\Lambda} \nu_j m_{ji}^\Lambda\right)(\exp(\alpha_i) - 1) \\ & + \sum_{i\in\Lambda} \mu_i \left(\sum_{j\in\Lambda} m_{ij}^\Lambda \exp(\alpha_j - \alpha_i) + m_{i0}^\Lambda \exp(-\alpha_i) - 1\right), \end{aligned} \tag{10.3}$$

and

$$\nu_j^\Lambda(\alpha) = \left(\nu_j + \sum_{i\in\Lambda}(\mu_i \exp(-\alpha_i) - \nu_i) m_{ij}^\Lambda\right) \exp(\alpha_j), \qquad j \in \Lambda^c, \tag{10.4}$$

where

$$\alpha_j = \log\left(\sum_{j\in\Lambda^c} m^{\Lambda}_{jk}\exp(\alpha_k) + m^{\Lambda}_{j0}\right),$$

(ν_i) *is the solution of the traffic equations* (1.2) *and*

$$m^{\Lambda}_{ij} = p_{ij} + \sum_{n=1}^{N}\sum_{k_1,\dots,k_n\in\Lambda^c} p_{ik_1}p_{k_1k_2}\cdots p_{k_nj},\ i\in\{1,\dots,N\},\ j\in\Lambda\cup\{0\}.$$

PROOF. Indeed, for $\beta\in\mathbb{R}^N$ with $\beta_\Lambda = \alpha_\Lambda$ and $\beta_{\Lambda^c} = 0$, the relation (8.6) gives

$$\begin{aligned}R(\alpha(\beta)) &= \sum_{i\in\Lambda}\lambda_i(\exp(\alpha_i)-1) + \sum_{j\in\Lambda^c}\lambda_j(\exp(\alpha_j(\beta))-1)\\ &\quad + \sum_{i\in\Lambda}\mu_i\Biggl(\sum_{j\in\Lambda} p_{ij}\exp(\alpha_j-\alpha_i) + p_{i0}\exp(-\alpha_i)\\ &\qquad + \sum_{j\in\Lambda^c} p_{ij}\exp(\alpha_j(\beta)-\alpha_i) - 1\Biggr)\end{aligned}\tag{10.5}$$

and the identity (8.2) implies that

$$\exp(\alpha_j(\beta)) = \sum_{k\in\Lambda} m^{\Lambda}_{jk}\exp(\alpha_k) + m^{\Lambda}_{j0}, \qquad j\in\Lambda^c.\tag{10.6}$$

It is clear that for every $j\in\Lambda^c$,

$$\sum_{k\in\Lambda} m^{\Lambda}_{jk} + m^{\Lambda}_{j0} = 1,$$

and for all $i,k\in\Lambda$,

$$p_{ik} + \sum_{j\in\Lambda^c} p_{ij}m^{\Lambda}_{jk} = m^{\Lambda}_{ik}$$

and hence, using (10.5) together with (10.6) and (10.1) we get that for any $\alpha\in\mathscr{D}_\Lambda$, $R(\alpha) = H_\Lambda(\alpha_\Lambda)$ where

$$\begin{aligned}H_\Lambda(\alpha_\Lambda) &= \sum_{i\in\Lambda}\left(\lambda_i + \sum_{j\in\Lambda^c}\lambda_j m^{\Lambda}_{ji}\right)(\exp(\alpha_i)-1)\\ &\quad + \sum_{i\in\Lambda}\mu_i\left(\sum_{j\in\Lambda} m^{\Lambda}_{ij}\exp(\alpha_j-\alpha_i) + m^{\Lambda}_{i0}\exp(-\alpha_i) - 1\right).\end{aligned}\tag{10.7}$$

Furthermore, by iterating the traffic equations

$$\nu_i = \lambda_i + \sum_{j\in\Lambda}\nu_j p_{ji} + \sum_{j\in\Lambda^c}\nu_j p_{ji}, \qquad i = 1,\dots,N$$

(where we will iterate only the third term in the right-hand side) we obtain

$$\nu_i = \lambda_i + \sum_{j\in\Lambda} \nu_j m^{\Lambda}_{ji} + \sum_{j\in\Lambda^c} \lambda_j m^{\Lambda}_{ji}, \qquad i = 1, \dots, N \tag{10.8}$$

and using the last relation together with (10.7) we get (10.3).

To prove the equality (10.4), we notice that for $\beta \in \mathbb{R}^N$ with $\beta_\Lambda = \alpha_\Lambda$ and $\beta_{\Lambda^c} = 0$, the identity (7.9) implies

$$\nu^{\Lambda}_j(\alpha(\beta)) = \left(\lambda_j + \sum_{i\in\Lambda^c} \lambda_i m^{\Lambda}_{ij} + \sum_{i\in\Lambda} \mu_i m^{\Lambda}_{ij} \exp(-\alpha_i)\right) \exp(\alpha_j(\beta)), \qquad j \in \Lambda^c$$

and hence, using again relations (10.6) and (10.8) we obtain (10.4). Lemma 10.1 is therefore proved. □

Consider now the function

$$H^*_\Lambda(v) = \sup_{\alpha\in\mathscr{D}_\Lambda} \{\langle\alpha, v\rangle - R(\alpha)\}. \tag{10.9}$$

The following proposition proves the first part of Theorem 2.

PROPOSITION 10.1. *For each $\Lambda \subseteq \{1, \dots, N\}$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, the supremum* (10.9) *is achieved in the unique point $\tilde\alpha^v \in D_\Lambda$; $\tilde\alpha^v_\Lambda$ is a unique solution of the system*

$$\nabla H_\Lambda(\alpha_\Lambda) = v_\Lambda \tag{10.10}$$

and for $i \in \Lambda^c$,

$$\tilde\alpha^v_i = \log\left(\sum_{j\in\Lambda} m^{\Lambda}_{ij} \exp(\tilde\alpha^v_j) + m^{\Lambda}_{i0}\right). \tag{10.11}$$

PROOF. Indeed, let $\alpha(\beta)$ be the unique solution of the system (8.1) for given Λ. Proposition 8.1 proves that:

(i) The mapping $\beta \to \alpha(\beta)$ defines a homeomorphism from the convex set $\{\beta: \beta_{\Lambda^c} = 0\}$ to $\mathscr{D}_\Lambda$.
(ii) The function $R(\alpha(\cdot))$ is strictly convex on $\{\beta: \beta_{\Lambda^c} = 0\}$.
(iii) The set

$$\{\beta: R(\alpha(\beta)) \le 0 \text{ and } \beta_{\Lambda^c} = 0\}$$

is compact and nonempty because $R(\alpha(0)) = R(0) = 0$.

This implies that the supremum (10.9) is achieved in a unique point. Denote this point by $\tilde\alpha^v$, then the relations (10.11) hold because $\tilde\alpha^v \in \mathscr{D}_\Lambda$. Moreover, in view of Lemma 10.1,

$$H^*_\Lambda(v) = \sup_{\alpha_\Lambda\in\mathbb{R}^\Lambda} \{\langle\alpha_\Lambda, v_\Lambda\rangle - H_\Lambda(\alpha_\Lambda)\}$$

and clearly, $\tilde{\alpha}^v_\Lambda$ is a unique point which achieves the supremum in the right-hand side of the above equality. This implies that $\tilde{\alpha}^v_\Lambda$ is a unique solution of the equation $\nabla H_\Lambda(\alpha_\Lambda) = v_\Lambda$ and Proposition 10.1 is therefore proved. □

Clearly, $\mathscr{D}_\Lambda \subset \mathscr{B}_\Lambda$ and hence, for any $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$,

$$\begin{aligned} H^*_\Lambda(v_\Lambda) &= \sup_{\alpha \in D_\Lambda} \{\langle \alpha, v\rangle - R(\alpha)\} \\ &\le \sup_{\alpha \in \mathscr{B}_\Lambda} \{\langle \alpha, v\rangle - R(\alpha)\} = l^\Lambda(v). \end{aligned} \tag{10.12}$$

The following statement gives a necessary and sufficient condition for

$$l^\Lambda(v) = H^*_\Lambda(v_\Lambda)$$

and completes the proof of Theorem 2. □

PROPOSITION 10.2. *Let $\Lambda \subseteq \{1, \dots, N\}$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$. Consider $\tilde{\alpha}^v \in \mathscr{D}_\Lambda$ for which* (10.10) *holds. Then*:

(i) *$l^\Lambda(v) = H^*_\Lambda(v_\Lambda)$ if*

$$\nu^\Lambda_j(\tilde{\alpha}^v) \le \mu_j \quad \text{for all } j \in \Lambda^c. \tag{10.13}$$

(ii) *Suppose that* (10.13) *does not hold, and consider*

$$\widetilde{\Lambda} = \{j \in \Lambda^c : \nu^\Lambda_j(\tilde{\alpha}^v) > \mu_j\},$$

then

$$l^\Lambda(v) = \min_{j \in \widetilde{\Lambda}} l^{\Lambda \cup \{j\}}(v). \tag{10.14}$$

PROOF. Consider the point $\alpha^v \in \mathscr{B}_\Lambda$ which achieves the maximum in the right-hand side of (10.12). Because of Proposition 8.3, the above $\alpha^v \in \mathscr{B}_\Lambda$ is unique and hence $l^\Lambda(v) = H^*_\Lambda(v_\Lambda)$ if and only if $\alpha^v = \tilde{\alpha}^v$.

Furthermore, let $\alpha(\cdot)$ be the unique solution of the system (8.1). Recall that $\alpha^v = \alpha(\beta^v)$ where $\beta^v \in \mathbb{R}^{\Lambda, N}_{\le 0}$ is a unique point which achieves the maximum of the strictly concave function $\langle \beta, v\rangle - R(\alpha(\beta))$ in $\mathbb{R}^{\Lambda, N}_{\le 0}$ (see Proposition 8.3), and consider $\tilde{\beta}^v \in \mathbb{R}^{\Lambda, N}_{\le 0}$ such that $\tilde{\beta}^v_\Lambda = \tilde{\alpha}^v_\Lambda$ and $\tilde{\beta}^v_{\Lambda^c} = 0$. Then obviously $\tilde{\alpha}^v = \alpha(\tilde{\beta}^v)$ and hence, $\alpha^v = \tilde{\alpha}^v$ if and only if $\beta^v = \tilde{\beta}^v$. It is clear that the above equality is verified if and only if

$$\nabla_{\beta_\Lambda} R(\alpha(\beta))|_{\beta = \tilde{\beta}^v} = v_\Lambda \tag{10.15}$$

and

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta))\big|_{\beta = \tilde{\beta}^v} \le 0. \tag{10.16}$$

Since

$$\nabla_{\beta_\Lambda} R(\alpha(\beta))\big|_{\beta = \tilde{\beta}^v} = \nabla H_\Lambda(\alpha_\Lambda)\big|_{\alpha_\Lambda = \tilde{\alpha}^v_\Lambda},$$

then (10.15) holds because of Proposition 10.1. Moreover, Proposition 8.2 shows that

$$\nabla_{\beta_{\Lambda^c}} R(\alpha(\beta))\big|_{\beta=\tilde{\beta}^v} = \nu^{\Lambda}_{\Lambda^c}(\alpha(\tilde{\beta}^v)) - \mu_{\Lambda^c}(\alpha(\tilde{\beta}^v)) = \nu^{\Lambda}_{\Lambda^c}(\tilde{\alpha}^v) - \mu_{\Lambda^c}(\tilde{\alpha}^v),$$

where the vectors $\nu^{\Lambda}_{\Lambda^c}(\tilde{\alpha}^v)$ and $\mu_{\Lambda^c}(\tilde{\alpha}^v)$ are defined by (7.9) and (7.6), respectively. Since $\tilde{\alpha}^v \in \mathscr{D}_{\Lambda}$, then (7.6) gives

$$\mu_j(\tilde{\alpha}^v) = \mu_j\left(\sum_{k=1}^{N} p_{jk}\exp(\tilde{\alpha}^v_k - \tilde{\alpha}^v_j) + p_{j0}\exp(-\tilde{\alpha}^v_j)\right) = \mu_j \quad \text{for all } j \in \Lambda^c$$

and therefore (10.16) is equivalent to (10.13). Hence $\beta^v = \tilde{\beta}^v$ if and only if (10.13) holds. This proves the first part of our proposition.

Suppose now that there exists $i \in \Lambda^c$ such that $\nu^{\Lambda}_i(\tilde{\alpha}^v) > \mu_i$ and let

$$\widetilde{\Lambda} = \{i \in \Lambda^c : \nu^{\Lambda}_i(\tilde{\alpha}^v) > \mu_i\}.$$

Consider again $\tilde{\beta}^v \in \mathbb{R}^N$ such that $\tilde{\beta}^v_{\Lambda} = \tilde{\alpha}^v_{\Lambda}$ and $\tilde{\beta}^v_{\Lambda^c} = 0$. Then

$$\frac{\partial}{\partial\beta_i} R(\alpha(\beta))\bigg|_{\beta=\tilde{\beta}^v} = \nu^{\Lambda}_i(\tilde{\alpha}^v) - \mu_i > 0 \quad \text{for all } i \in \widetilde{\Lambda},$$

$$\frac{\partial}{\partial\beta_i} R(\alpha(\beta))\bigg|_{\beta=\tilde{\beta}^v} = \nu^{\Lambda}_i(\tilde{\alpha}^v) - \mu_i = 0 \quad \text{for all } i \in \Lambda^c\setminus\widetilde{\Lambda}$$

and

$$\frac{\partial}{\partial\beta_i} R(\alpha(\beta))\bigg|_{\beta=\tilde{\beta}^v} = V^{\Lambda}_i(\tilde{\alpha}^v) = v_i \quad \text{for all } i \in \Lambda.$$

Because the function $R(\alpha(\cdot))$ is strictly convex on $\mathbb{R}^{\Lambda, N}_{\le 0}$, the above relations imply that $\tilde{\beta}^v$ achieves the maximum of the function $\beta \to \langle \beta, v\rangle - R(\alpha(\beta))$ in the subset

$$\left\{\beta \in \mathbb{R}^{\Lambda, N}_{\le 0} : \beta_{\widetilde{\Lambda}} = 0\right\} \subset \mathbb{R}^{\Lambda, N}_{\le 0}$$

and hence, $\tilde{\alpha}^v = \alpha(\tilde{\beta}^v)$ achieves the maximum of the function $\alpha \to \langle \alpha, v\rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda} \cap \mathscr{D}_{\Lambda}$.

Consider now the point α^v which achieves the maximum of the function $\langle \alpha, v\rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda}$. It is clear that in this case $\alpha^v \ne \tilde{\alpha}^v$ and since $\tilde{\alpha}^v$ achieves the maximum of the function $\alpha \to \langle \alpha, v\rangle - R(\alpha)$ in $\mathscr{B}_{\Lambda} \cap \mathscr{D}_{\Lambda}$, then $\alpha^v \notin D_{\widetilde{\Lambda}}$ and hence there exists $i \in \widetilde{\Lambda}$ such that

$$\alpha^v_i < \log\left(\sum_{j=1}^{N} p_{ij}\exp(\alpha^v_j) + p_{i0}\right).$$

Using now the last part of Proposition 8.3 we conclude that there exists $i \in \widetilde{\Lambda}$ such that

$$l^{\Lambda}(v) = l^{\Lambda \cup \{i\}}(v). \tag{10.17}$$

But

$$\begin{aligned} l^{\Lambda}(v) &= \sup_{\alpha \in \mathscr{B}_{\Lambda}} \{\langle \alpha, v \rangle - R(\alpha)\} \le \min_{i \in \widetilde{\Lambda}} \sup_{\alpha \in \mathscr{B}_{\Lambda \cup \{i\}}} \{\langle \alpha, v \rangle - R(\alpha)\} \\ &= \min_{i \in \widetilde{\Lambda}} l^{\Lambda \cup \{i\}}(v) \end{aligned}$$

because $\mathscr{B}_{\Lambda} \subset \mathscr{B}_{\Lambda \cup \{i\}}$ for all $i \in \widetilde{\Lambda}$, and hence relation (10.17) yields (10.14). Proposition 10.2 is proved. □

APPENDIX

In this section we prove Proposition 8.1. For this, we shall use the following lemma.

LEMMA A.1. *Consider the function*

$$r(\alpha) = \sum_{y \in \mathbb{Z}^N} a(y) e^{\langle \alpha, y \rangle}, \qquad \alpha \in \mathbb{R}^N, \tag{A.1}$$

where $a(y) \ge 0$ *for all* $y \in \mathbb{Z}^N$, *and let the set* $\mathscr{O} = \{y \in \mathbb{Z}^N : a(y) \ne 0\}$ *be finite.*

(i) *Suppose that the set* $\mathscr{O}$ *contains a basis in* $\mathbb{R}^N$. *Then the function* $r(\cdot)$ *is strictly convex everywhere on* $\mathbb{R}^N$.

(ii) *Suppose moreover that for any* $y \in \mathbb{Z}^N$ *there exists* $n \in \mathbb{N}$ *and there exist* $y_0, \dots, y_n \in \mathscr{O}$ *such that* $y = y_0 + \cdots + y_n$. *Then for any* $v \in \mathbb{R}^N$ *and for any* $c \in \mathbb{R}_+$ *the level set*

$$\{\alpha \in \mathbb{R}^N : r(\alpha) - \langle \alpha, v \rangle \le c\}$$

is a compact subset of $\mathbb{R}^N$.

PROOF. Indeed, using (A.1) it follows that for all $\alpha, \xi \in R^N$,

$$\langle \xi, \partial_{\alpha}^2 r(\alpha) \xi \rangle = \sum_{i,j=1}^{N} \frac{\partial^2}{\partial \alpha_i \partial \alpha_j} r(\alpha) \xi_i \xi_j = \sum_{y \in \mathbb{Z}^N} a(y) e^{\langle \alpha, y \rangle} \langle \xi, y \rangle^2. \tag{A.2}$$

Suppose that $\mathscr{O}$ contains a basis in $\mathbb{R}^N$, then for any $\xi \in \mathbb{R}^N$, there exists $y \in \mathscr{O}$ such that $\langle \xi, y \rangle \ne 0$ and hence, using (A.2), we get

$$\langle \xi, \partial_{\alpha}^2 r(\alpha) \xi \rangle > 0 \quad \text{ for all } \alpha, \xi \in \mathbb{R}^N.$$

The above implies that the function $r(\cdot)$ is strictly convex, and so the first part of our lemma is proved.

Let us prove the second one. Since the function $\alpha \to r(\alpha) - \langle \alpha, v\rangle$ is continuous for any $v \in \mathbb{R}^N$, it is sufficient to show that the set

$$\{\alpha \in \mathbb{R}^N\colon r(\alpha) - \langle \alpha, v\rangle \le c\}$$

is bounded for all $v \in \mathbb{R}^N$ and $c \in \mathbb{R}_+$.

Suppose that for any $y \in \mathbb{Z}^N$, there exists $n \in \mathbb{N}$ and there exist $y_0, \dots, y_n \in \mathscr{O}$ such that $y = y_0 + \cdots + y_n$. Then for any $v \in \mathbb{R}^N$, there exist $v_y \ge 0,\ y \in \mathscr{O}$ such that

$$v = \sum_{y\in\mathscr{O}} v_y y,$$

and therefore,

$$\begin{aligned}\langle \alpha, v\rangle - r(\alpha) &= \sum_{y\in\mathscr{O}} (v_y\langle \alpha, y\rangle - a(y)\exp(\langle \alpha, y\rangle))\\ &\le \sum_{y\in\mathscr{O}} \sup_{t\in\mathbb{R}}(v_y t - a(y)e^t) \le \sum_{y\in\mathscr{O}} (v_y \log(v_y/a(y)) - v_y).\end{aligned}$$

The above implies that the Fenchel–Legendre transform of the function $r(\cdot)$,

$$r^*(v) = \sup_{\alpha\in\mathbb{R}^N} (\langle \alpha, v\rangle - r(\alpha))$$

is finite and continuous everywhere in $\mathbb{R}^N$ (see [24]). Finally, using

$$\langle \alpha, v\rangle \le r(\alpha) + r^*(v), \qquad \alpha, v \in \mathbb{R}^N,$$

we get

$$\sup_{v'\in\mathbb{R}^N\colon |v'|\le 1} \langle \alpha, v'\rangle \le c + \sup_{v'\in\mathbb{R}^N\colon |v'|\le 1} r^*(v+v')$$

for all $\alpha \in \mathbb{R}^N$ such that $r(\alpha) - \langle \alpha, v\rangle \le c$, and we conclude therefore, that the set $\{\alpha \in \mathbb{R}^N\colon r(\alpha) - \langle \alpha, v\rangle \le c\}$ is bounded for all $v \in \mathbb{R}^N$ and $c \in \mathbb{R}_+$. Thus the second part of our lemma is also proved. □

Because of Assumption A, the function $R(\cdot)$ clearly satisfies the conditions of Lemma A.1 and hence, using this lemma, we immediately get the following statement.

COROLLARY A.1. *The function $R(\cdot)$ is strictly convex and the set*

$$\{\alpha \in \mathbb{R}^N\colon R(\alpha) - \langle \alpha, v\rangle \le c\}$$

is a compact subset of $\mathbb{R}^N$ for all $v \in \mathbb{R}^N$ and $c \in \mathbb{R}_+$.

We are ready now to prove our proposition.

PROOF OF PROPOSITION 8.1. Indeed, the system (8.1) is equivalent to the following one:

$$\text{(A.3)}\qquad \begin{cases} \alpha_i = \beta_i, & \text{for } i \in \Lambda, \\ \exp(\alpha_i) = \sum_{j\in\Lambda^c} p_{ij}\exp(\beta_i)\exp(\alpha_j) & \\ \qquad\qquad + \sum_{j\in\Lambda} p_{ij}\exp(\beta_i+\beta_j) + p_{i0}\exp(\beta_i), & \text{for } i \in \Lambda^c. \end{cases}$$

Moreover, because of Assumption A, the matrix $(p_{ij})_{i,\, j\in\Lambda^c}$ has a spectrum radius less then unity and hence there exists $\varepsilon_0 > 0$ such that for $\beta \in \mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$, the matrix $(p_{ij}e^{\beta_i})_{i,\, j\in\Lambda^c}$ has also a spectrum radius less then unity. The above implies that for $\beta \in \mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$, the system (A.3) has a unique solution and the iterating method applied to (A.3) gives (8.2). The first part of our lemma is therefore proved.

Let us prove the second one. Indeed, the implicit function theorem applied to the system (8.1) implies that $\alpha(\cdot) \in C^\infty$ everywhere in $\mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$, and therefore $R(\alpha(\cdot)) \in C^\infty$ also everywhere in $\mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$.

Let us verify that the function $R(\alpha(\cdot))$ is strictly convex everywhere on $\mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$. Indeed, observe that for any $\beta \in \mathbb{R}^{\Lambda,N}_{<\varepsilon_0}$,

$$\text{(A.4)}\qquad \begin{aligned} R(\alpha(\beta)) = {}& \sum_{j\in\Lambda^c} \mu_j(\exp(-\beta_j)-1) + \sum_{j\in\Lambda} \lambda_j(\exp(\beta_j)-1) \\ & + \sum_{j\in\Lambda^c} \lambda_j(\exp(\alpha_j(\beta))-1) \\ & + \sum_{j\in\Lambda} \mu_j\Bigg(\sum_{k\in\Lambda} p_{jk}\exp(\beta_k-\beta_j) + p_{j0}\exp(-\beta_j) \\ & \qquad\qquad + \sum_{k\in\Lambda^c} p_{jk}\exp(\alpha_k(\beta)-\beta_j) - 1\Bigg) \end{aligned}$$

and consider the functions

$$\text{(A.5)}\qquad \begin{aligned} r_1(\beta) = {}& \sum_{j\in\Lambda^c} \mu_j(\exp(-\beta_j)-1) \\ & + \sum_{j\in\Lambda} \mu_j\Bigg(\sum_{k\in\Lambda} p_{jk}\exp(\beta_k-\beta_j) + p_{j0}\exp(-\beta_j) - 1\Bigg) \end{aligned}$$

and

$$\text{(A.6)}\qquad r_2(\alpha) = \sum_{j=1}^{N} \lambda_j(\exp(\alpha_j)-1) + \sum_{j\in\Lambda,\, k\in\Lambda^c} \mu_j p_{jk}\exp(\alpha_k-\alpha_j).$$

Then $R(\alpha(\cdot)) = r_1(\cdot) + r_2(\alpha(\cdot))$ and hence to prove that the function $R(\alpha(\cdot))$ is strictly convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ it is sufficient to show that the function $r_1(\cdot)$ is strictly convex and the function $r_2(\alpha(\cdot))$ is convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$.

To verify that the function $r_2(\alpha(\cdot))$ is convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$, we have to show that for all $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ and $\xi \in \mathbb{R}^N$,

$$\langle \xi, \partial^2_\beta r_2(\alpha(\beta))\xi\rangle = \sum_{i, j=1}^{N} \frac{\partial}{\partial\beta_i \partial\beta_j} r_2(\alpha(\beta))\xi_i\xi_j \geq 0. \tag{A.7}$$

For this we notice that the function $r_2(\cdot)$ is convex everywhere on $\mathbb{R}^N$ and for any $i \in \Lambda^c$ the function $\alpha_i(\cdot)$ is convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ as a limit of convex functions [see (8.2)]. The above implies that for all $\beta \in \mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ and $\xi \in \mathbb{R}^N$,

$$\begin{aligned}\langle \xi, \partial^2_\beta r_2(\alpha(\beta))\xi\rangle &= \langle \xi(\beta), \partial^2_\alpha r_2(\alpha)\xi(\beta)\rangle \\ &\quad + \sum_{j\in\Lambda^c} \frac{\partial}{\partial\alpha_j} r_2(\alpha)\bigg|_{\alpha=\alpha(\beta)} \langle \xi, \partial^2_\beta \alpha_j(\beta)\xi\rangle \geq 0,\end{aligned}$$

where

$$\xi_j(\beta) = \sum_{i=1}^{N} \frac{\partial}{\partial\beta_i}\alpha_j(\beta)\xi_i, \qquad j = 1, \dots, N$$

and

$$\langle \xi(\beta), \partial^2_\alpha r_2(\alpha)\xi(\beta)\rangle = \sum_{i, j=1}^{N} \frac{\partial^2}{\partial\alpha_i\partial\alpha_j} r_2(\alpha)\xi_i(\alpha)\xi_j(\alpha) \geq 0$$

because the function $r_2(\cdot)$ is convex everywhere on $\mathbb{R}^N$,

$$\langle \xi, \partial^2_\beta \alpha_i(\beta)\xi\rangle = \sum_{j, j=k}^{N} \frac{\partial^2}{\partial\beta_j\partial\beta_k}\alpha_j(\beta)\xi_j\xi_k \geq 0 \quad \text{for all } i \in \Lambda^c,$$

because the functions $\alpha_i(\cdot)$, $i \in \Lambda^c$ are convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$ and

$$\frac{\partial}{\partial\alpha_j} r_2(\alpha) \geq 0 \quad \text{for all } j \in \Lambda^c \text{ and } \alpha \in \mathbb{R}^N.$$

Thus (A.7) is verified and hence, the function $r_2(\alpha(\cdot))$ is convex everywhere on $\mathbb{R}^{\Lambda, N}_{<\varepsilon_0}$.

Let us prove now that the function $r_1(\cdot)$ is strictly convex. In view of Lemma A.1, it is sufficient to show that for every $i \in \{1, \dots, N\}$ the ith unit vector ε^i is included in the linear space spanned by the set

$$\mathscr{O} = \{\varepsilon^i : i \in \Lambda^c\} \cup \{\varepsilon^i - \varepsilon^j : i, j \in \Lambda,\ p_{ji} \neq 0\} \cup \{\varepsilon^i : i \in \Lambda,\ p_{i0} \neq 0\}$$

(recall that by assumption $\mu_i > 0$ for all $i \in \{1, \dots, N\}$).

The above obviously holds for $i \in \Lambda^c$. Consider now ε^i with $i \in \Lambda$. Then because of Assumption (A) there exist a sequence $i_1, \dots, i_n \in \Lambda$ such that either

$$p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n} p_{i_n 0} > 0, \tag{A.8}$$

or for some $j \in \Lambda^c$,

$$p_{ii_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n} p_{i_n j} > 0. \tag{A.9}$$

If (A.8) holds, then $\varepsilon^{i_1} - e^i$, $\varepsilon^{i_2} - e^{i_1}$, $\varepsilon^{i_n} - \varepsilon^{i_{n-1}}$, $\varepsilon^{i_n} \in \mathscr{O}$, and hence, the vector

$$\varepsilon^i = -(\varepsilon^{i_1} - e^i + \varepsilon^{i_2} - e^{i_1} + \dots + \varepsilon^{i_n} - \varepsilon^{i_{n-1}}) + \varepsilon^{i_n}$$

is included to the linear space spanned by $\mathscr{O}$.

If (A.8) does not hold but (A.9) is verified, then $\varepsilon^{i_1} - \varepsilon^i$, $\varepsilon^{i_2} - \varepsilon^{i_1}$, $\varepsilon^{i_n} - \varepsilon^{i_{n-1}}$, $\varepsilon^j - \varepsilon^{i_n} \in \mathscr{O}$. Since $\varepsilon^j \in \mathscr{O}$ for $j \in \Lambda^c$, the above implies that in this case the vector

$$\varepsilon^i = -(\varepsilon^{i_1} - e^i + \varepsilon^{i_2} - e^{i_1} + \dots + \varepsilon^{i_n} - \varepsilon^{i_{n-1}} + \varepsilon^j - \varepsilon^{i_n}) + \varepsilon^j$$

is also included to the linear space spanned by $\mathscr{O}$.

Thus, the function $R(\alpha(\cdot))$ is strictly convex everywhere on $\mathbb{R}^{\Lambda, N}_{\varepsilon_0}$.

To complete the proof of our proposition, let us observe that the mapping $\alpha(\cdot)$ is an homeomorphism from $\mathbb{R}^{\Lambda, N}_{\le 0}$ to $\mathscr{B}_\Lambda$ and because of $v_{\Lambda^c} = 0$ the following equality holds:

$$\langle \alpha(\beta), v \rangle = \langle \beta, v \rangle.$$

The above immediately implies (8.3). To verify that for every $c \in \mathbb{R}_+$ and $v \in \mathbb{R}^N$ such that $v_{\Lambda^c} = 0$, the set $\{\beta \in \mathbb{R}^{\Lambda, N}_{\le 0} : \langle \beta, v \rangle - R(\alpha(\beta)) \le c\}$ is compact, it is sufficient to notice now that $\alpha(\cdot)$ is an homeomorphism from the above set to

$$\{\alpha \in \mathscr{B}_\Lambda : \langle \alpha, v \rangle - R(\alpha) \le c\},$$

which is compact because the set $\{\alpha \in \mathbb{R}^N : \langle \alpha, v \rangle - R(\alpha) \le c\}$ is compact for all $c \in \mathbb{R}_+$ and $v \in \mathbb{R}^N$ (see Corollary A.1) and the set $\mathscr{B}_\Lambda$ is closed. Proposition 8.1 is therefore proved.

REFERENCES

[1] ALANYALI, M. and HAJEK, B. (1998). On large deviations of Markov processes with discontinuous statistics. *Ann. Appl. Probab.* **8** 45–66.

[2] ATAR, R. and DUPUIS, P. (1999). Large deviations and queueing networks: methods for rate function identification. *Stochastic Process. Appl.* **84** 255–296.

[3] BLINOVSKI, V. M. and DOBRUSHIN, R. L. (1994). Process level large deviations for a class of piecewise homogeneous random walks. *Progr. Probab*. **34** 1–59.

[4] BOROVKOV, A. A. (1986). Limit theorems for queueing networks. *I. Teor. Veroyatnost. i Primenen* **31** 474–490.

[5] BOTVICH, D. D. and ZAMYATIN, A. A. (1995). On fluid approximations for conservative networks. *Markov Processes and Related Fields* **1** 113–140.

[6] CHEN, H. and MANDELBAUM, A. (1991). Discrete flow networks: bottleneck analysis and fluid approximations. *Math. Oper. Res*. **16** 408–446.

[7] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York.

[8] DOBRUSHIN, R. L. and PECHERSKY, E. A. (1994). Large deviations for tandem queueing systems. *J. Appl. Math. Stochastic Anal*. **7** 301–330.

[9] DUPUIS, P. and ELLIS, R. S. (1992). Large deviations for Markov processes with discontinuous statistics. II. Random walks. *Probab. Theory Related Fields* **91** 153–194.

[10] DUPUIS, P. and ELLIS, R. S. (1995). The large deviation principle for a general class of queueing systems. I. *Trans. Amer. Math. Soc*. **347** 2689–2751.

[11] DUPUIS, P. ELLIS, R. and WEISS, A. (1991). Large deviations for Markov processes with discontinuous statistics I: general upper bounds. *Ann. Probab*. **19** 1280–1297.

[12] DUPUIS, P. ISHII, H. and SONER, H. M. (1990). A viscosity solution approach to the asymptotic analysis of queueing systems. *Ann. Probab* **18** 226–255.

[13] DUPUIS, P. and RAMANAN, K. (1999). Convex duality and the Skorokhod problem. I. *Probab. Theory Related Fields* **115** 153–195.

[14] DURRETT, R. (1996). *Probability: Theory and Examples*, 2nd ed. Duxbury Press, Belmont, CA.

[15] FRATER, M. R., LENNON, T. M. and ANDERSON, B. D. O. (1991). Optimally efficient estimation of the statistics of rare events in queueing networks. *IEEE Trans. Automat. Control* **36** 1395–1405.

[16] FREIDLIN, M. I. and WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. Springer, New York.

[17] IGNATYUK, I. A., MALYSHEV, V. A. and SHCHERBAKOV, V. V. (1994). The influence of boundaries in problems on large deviations. *Uspekhi Mat. Nauk* **49** 43–102.

[18] ISCOE, I. and MCDONALD, D. (1995). Asymptotics of exit times for Markov jump processes. II: applications to Jackson networks. *Ann. Probab*. **22** 2168–2182.

[19] JACKSON, J. R. (1957). Networks of waiting lines. *Oper. Res*. **5** 518–521.

[20] KELLY, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley, Chichester.

[21] LABRÈCHE, N. and MCDONALD, D. (1995). Large deviations of the total backlog in a queueing network. Preprint.

[22] MALYSHEV, V. A. and MEN'SHIKOV, M. V. (1981). Ergodicity, continuity, and analyticity of countable Markov chains. *Trans. Moscow Math. Soc*. **1** 1–47.

[23] PAREKH, S. and WALRAND, J. (1989). A quick simulation method for excessive backlogs in networks of queues. *IEEE Trans. Automat. Control* **34** 54–66.

[24] ROCKAFELLAR, R. T. (1997). *Convex Analysis*, Princeton Univ. Press.

[25] ROGERS, L. C. G. and WILLIAMS, D. (1987). *Diffusions, Markov Processes, and Martingales* **2**: *Itô Calculus*. Wiley, New York.

[26] RUGET, G. (1979). Quelques occurences des grands écarts dans la littérature électronique. *Astérisque* **68** 187–199.

[27] SHARPE, M. (1988). *General Theory of Markov Processes*. Academic Press, Boston.

[28] SHWARTZ, A. and WEISS, A. (1995). *Large Deviations for Performance Analysis*. Chapman and Hall, London.

[29] TSOUCAS, P. (1992). Rare events in series of queues. *J. Appl. Probab*. **29** 168–175.

UNIVERSITÉ DE CERGY-PONTOISE
DÉPARTEMENT DE MATHÉMATIQUES, 2
AVENUE ADOLPHE CHAUVIN
95302 CERGY-PONTOISE CEDEX
FRANCE
E-MAIL: Irina.Ignatiouk@math.u-cergy.fr