How large queue lengths build up in a Jackson network

Kurt Majewski
Siemens AG, Corporate Technology
81730 Munich, Germany

and

Kavita Ramanan
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213

March, 2008

Abstract

We identify the asymptotically most likely way in which large queues build up in stationary Jackson networks. We characterize the optimal time-dependent large deviations change of measure on the service, arrival and routing processes that leads to this rare event. The proof uses pathwise time-reversal arguments that may be of independent interest.

Key words. Large deviations, single class networks, Jackson networks, minimizing trajectories, time reversal, rate function, fluid limits.

AMS Subject Classification. Primary: 60K25, 60F10; Secondary: 60G10.

1 Introduction

1.1 Background and Motivation

Large deviations is an asymptotic theory that is useful for obtaining both quantitative and qualitative insight into the probabilities of rare events. The theory typically expresses the exponential decay rate (or rate function) of the probability of a rare event in terms of the solution to an associated variational problem. In such situations, the minimizer of the variational problem provides insight into the most likely behavior of the system that leads to the rare event [18]. In the context of queueing networks, this knowledge can be useful for designing mechanisms that minimize the probabilities of undesirable rare events such as buffer overflows [32]. Identification of the minimizer is also useful for importance sampling methods that estimate the actual probabilities of the rare events, and not just their log equivalents, through simulation (see, for example, [5, 10, 17, 20, 30] and references therein).

Queueing networks pose a significant challenge to large deviations theory due to the fact that they exhibit discontinuities in the transition rates at the boundaries of the state space. Despite some recent progress in extending the classical theory to handle discontinuities [11, 13, 22, 28], there are still many situations of practical interest where even the existence of a large deviation principle has not been shown, let alone the explicit form of the rate function identified or probabilities of specific rare events computed. A majority of the explicit results are confined to one or two dimensions [1, 14, 29, 31, 33]. Jackson networks are more tractable due to the fact that they have a known product-form stationary distribution [23, 24]. Despite this simplicity, an explicit expression for the local large deviation rate function for Jackson queueing networks was only recently obtained in [22]. The result in [22] yields expressions for the exponential decay rates of probabilities of a large class of rare events associated with Jackson networks in terms of variational problems involving the local rate function. However, the solution of these variational problems to identify the associated minimizing trajectories appears to be a rather difficult task in many cases. In particular, explicit identification of the minimizer of the variational problem that characterizes the tail probabilities of the queue lengths appears to be rather hard for dimensions greater than two.

In this work we overcome these difficulties in the special case of Jackson single server networks by combining two techniques. The first is to use the partial continuity of the relation between network primitives and the behavior of the underlying functional network model, which exists despite the non-uniqueness in this relation. In contrast to a direct application of large deviation principles for the Jackson queue length processes [22, 28], this allows us to retain the connection to the network primitive processes. The mentioned partial continuity was identified in [25], but only recently strengthened to cover fluid limits of Jackson networks in [27]. The model in [27] unifies discrete customer and fluid models, and therefore turns out to be particularly well-suited for connecting the Markov chain describing the Jackson network with the differential equations that capture the asymptotically most probable behavior, conditioned on large queue lengths at time 0, as needed in this work.

The second technique is time reversal. For a single station with Poisson arrivals at rate α and exponential service with rate $\mu > \alpha$, it is well known that the large deviations minimizing path to overflow is obtained by exchanging arrival and service rates [31]. In other words, the large deviation path to reach a certain queue level is the time-reversal of the path the fluid limit uses to drain to zero. In [31] a two-dimensional large deviations problem was reduced to a one-dimensional problem, and time-reversal ideas were used to identify minimizing large deviations paths. However, there have been relatively few rigorous results in the multi-dimensional setting [2]. In our work we define time reversal with respect to time 0 for fluid network primitives and network behavior on the doubly infinite time interval. This general construction allows us to exploit time reversal for the large deviation analysis in the general n-dimensional case.

The developed techniques may potentially be applied to more general multi-dimensional networks, e.g., Kelly or even generalized Jackson networks, for which there are relatively few results. Indeed, our proof is independent of dimension and, in contrast to [2], makes no explicit use of either the specific form of the stationary distribution or the Markov property of the queue length process in a Jackson network. Instead, we use a representation of the large deviations optimization problem that is obtained using a kind of contraction principle formulation, expressed in terms of the local rate function of the primitive processes [26], much in the spirit of [16], rather than the representation obtained in [22, 28], which is in terms of the local rate function of the queue length processes. This enables us to identify the most likely arrival, service and routing rates that lead to the build up a large queue, and not just the most likely path of the queue itself.

The heart of the proof lies in a reformulation of this large deviations optimization problem in terms of a "fluid limit" optimal control problem using *deterministic* time reversal arguments, which are somewhat subtle due to the presence of boundaries. The known exponential decay rate of the tails of the stationary queue length distribution falls out as a consequence of our analysis. We believe that this general philosophy is likely to have broader applicability in identifying minimizing trajectories for large deviation problems with boundaries. Indeed, deterministic time reversal arguments have also proved useful in the context of reflected diffusions in [16], where they were used to identify minimizing large deviation paths for skew-symmetric reflected Brownian motions. However, unlike here, the proof in [16] uses the explicit form of the stationary distribution in order to identify the minimizing trajectory, which is undesirable from the point of view of extending the technique to analyze more general networks for which the stationary distribution is not known.

The outline of the paper is as follows. In Section 2, we first present a model of the behavior of Jackson networks on the time interval $\mathbb{R}$, which includes cumulative arrival, service and routing processes. Section 3 contains a summary of our main results. Roughly speaking, our main result states that, conditioned on a given queue length at time 0, the sequence of scaled queueing processes converge "exponentially fast" to an asymptotic limit. An explicit description of the asymptotic limit is provided in Section 3.1 and a rigorous statement of the main result, Theorem 7, is provided in Section 3.2. An explicit example of the trajectory to overflow in a three station network is provided in Section 3.3. The proof of the main result relies on certain pathwise time-reversal arguments that may be of independent interest. Specifically, a functional description of a time-reversed network is provided in Section 4. Lastly, the large deviations analysis is carried out in Section 5, culminating in the proof of Theorem 7 in Section 5.5.

1.2 Notation and Terminology

We first introduce some function spaces that will be used throughout the paper. For $c \in \mathbb{R}$, we use $\mathcal{D}_c$ to denote the set of right continuous paths F from $\mathbb{R}$ to $\mathbb{R}$ that have finite left limits and satisfy $\lim_{|t|\to\infty} F(t)/t = c$. We let $\mathcal{I}_c$ denote the subset of non-decreasing paths in $\mathcal{D}_c$. Moreover, for vectors $v \in \mathbb{R}^n$ (respectively, matrices $M \in \mathbb{R}^{n\times n}$), we let $\mathcal{D}_v$ (respectively, $\mathcal{D}_M$) be the space of vectors (respectively, matrices) of paths whose ith (respectively, (i, j)th) component lies in $\mathcal{D}_{v_i}$ (resp. $\mathcal{D}_{M_{i,j}}$), and analogously define $\mathcal{I}_v$ (resp. $\mathcal{I}_M$). We equip these function spaces with the topology induced by the *quasi-linearly discounted uniform norm* defined by

$$\|D\| \quad := \quad \sup_{t \in \mathbb{R}} \frac{|D(t)|}{1+|t|}. \tag{1}$$

Product function spaces are equipped with the corresponding product topology.

Function spaces will be equipped with the σ-algebra generated by the family of one-dimensional projections. We will be interested in stochastic processes defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose paths lie in function spaces of the type described above. Given some function space $\mathcal{S}$, by some abuse of notation, we will sometimes use the notation $X \in \mathcal{S}$ to mean that the paths of the stochastic process X lie $\mathbb{P}$-a.s. in $\mathcal{S}$.

As a general convention, throughout the paper, we will use upper case F to denote paths and, when the path is absolutely continuous, the corresponding lower case $f = \dot{F}$ to denote a derivative of the path.

We abbreviate the set $\{1, \ldots, n\}$ by $\mathcal{J}$, denote the indicator function of a set M with 1_M, use $\lfloor c \rfloor$ to specify the largest element of $\mathbb{Z}$ smaller or equal to $c \in \mathbb{R}$, and let $\mathrm{diag}(v)$ be the diagonal matrix with diagonal entries taken from the components of the vector v. Lastly, we let $e \in \mathbb{R}^n$ be the vector having ones in each component and $I := \mathrm{diag}(e)$.

2 Jackson network model

We now introduce equations that describe the behavior of a single-server Jackson network with n nodes on the time interval $\mathbb{R}$. Let $\alpha \in \mathbb{R}^n_+$ denote the vector of mean exogenous arrival rates, let $\mu \in \mathbb{R}^n_+$ represent the vector of mean service rates, and let $\gamma \in \mathbb{R}^{n \times n}_+$ be the substochastic matrix that specifies the routing rates. For $j \in \mathcal{J}$, the value $\gamma_{0,j} \in \mathbb{R}_+$ defined by

$$\gamma_{0,j} \quad := \quad 1 - \sum_{i \in \mathcal{J}} \gamma_{i,j} \tag{2}$$

is the mean rate at which customers that have completed service at queue j leave the network. This setup is illustrated in Figure 1.

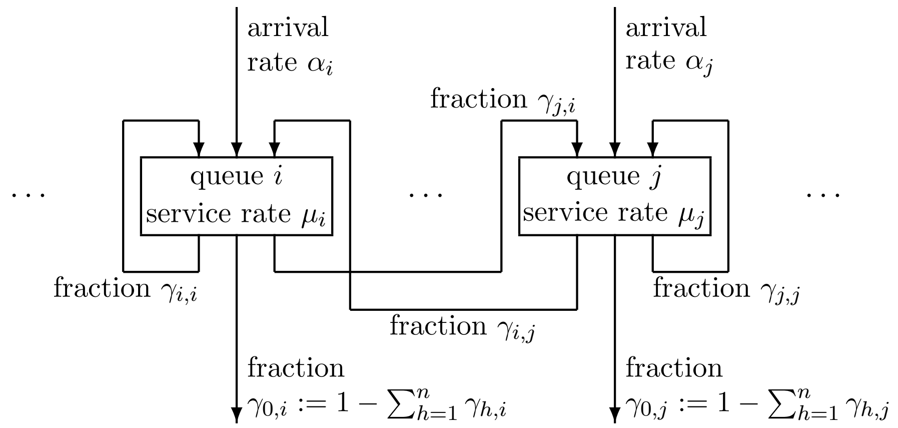

Figure 1: Jackson network

We assume that γ has spectral radius strictly less than 1. Hence the matrix $I - \gamma$ has a non-negative inverse and the equation

$$\lambda \;=\; \alpha + \gamma\lambda \tag{3}$$

has a unique solution given by $\lambda = (I - \gamma)^{-1}\alpha$, which defines the total arrival rates at the queues (see, for example, [6]). We assume throughout this work that all queues have non-zero total arrival rates and that the network is stable, that is, the condition

$$0 < \lambda < \mu$$

is satisfied. In particular, every component of μ is strictly positive and the vector of traffic intensities $\varrho \in \mathbb{R}^n_+$ defined by

$$\varrho_i \;:=\; \lambda_i/\mu_i \tag{4}$$

satisfies $0 < \varrho < e$, where e is the vector that has a one in every component.

We now define the basic primitive processes describing the Jackson network. For $i \in \mathcal{J}$ we let $\tilde{A}_i = (\tilde{A}_i(t))_{t\in\mathbb{R}}$ be a Poisson counting process on $\mathbb{R}$ with intensity α_i that satisfies $\tilde{A}_i(0) = 0$. The jump times of the process $\tilde{A}_i$ describe the arrival times of exogenous customers to queue i. Similarly, we let $\tilde{S}_i = (\tilde{S}_i(t))_{t\in\mathbb{R}}$ be a Poisson counting process on $\mathbb{R}$ with intensity σ_i that satisfies $\tilde{S}_i(0) = 0$. For $i \in \mathcal{J}$ and $j \in \mathbb{Z}$, the interval between the jth and $(j+1)$th jump times of the process $\tilde{S}_i$ specifies the service time of the jth customer at queue i. Moreover, let $(\hat{R}_{j,h})_{h\in\mathbb{Z}}$ be a sequence of i.i.d. $\{0,\ldots,n\}$-valued random variables with $\mathbb{P}(\hat{R}_{j,0}=i)=\gamma_{i,j}$ describing the routing decisions of node j. We let $\tilde{R}_j=(\tilde{R}_{1,j},\ldots,\tilde{R}_{n,j})^T$ be the n-dimensional partial sums process on $\mathbb{R}$, in which for $i\in\mathcal{J}$, the ith component $(\tilde{R}_{i,j}(t))_{t\in\mathbb{R}}$ is defined by

$$\tilde{R}_{i,j}(t) \;:=\; \begin{cases} \sum_{h=1}^{\lfloor t\rfloor} 1_{\{\hat{R}_{j,h}=i\}}(h) & \text{if } t\geq 1,\\ 0 & \text{if } 0\leq t<1,\\ -\sum_{h=\lfloor t\rfloor+1}^{0} 1_{\{\hat{R}_{j,h}=i\}}(h) & \text{if } t<0, \end{cases}$$

where 1_M is the indicator function of the set M. The process $\tilde{R}_{i,j}$ describes the cumulative number of customers routed from queue j to queue i. We assume that the processes $\tilde{A}_1,\ldots,\tilde{A}_n$, $\tilde{S}_1,\ldots,\tilde{S}_n$, $\tilde{R}_1,\ldots,\tilde{R}_n$ are independent. From the definition it immediately follows that the triple of network primitives $(\tilde{A},\tilde{S},\tilde{R})$ lies in $\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma$ (recall the function space terminology given in Section 1.2).

The queue lengths process $\tilde{Q}$ associated with the network primitives is best described in terms of certain auxiliary componentwise non-decreasing network processes, $\tilde{X}$, $\tilde{Z}$, $\tilde{Y}$ and $\tilde{B}$ that represent, respectively, the cumulative number of arrivals to, the cumulative number of departures from, the cumulative idle time at and the cumulative busy time at the different nodes of the Jackson network as a function of time. In order to describe these processes, we will find it convenient to first introduce the following, more general, correspondence between network primitives (A,S,R) and the corresponding network behavior (X,Z,Q,Y,B).

Definition 1 (Network Equations) *The quintuple of paths (X,Z,Q,Y,B) are said to solve the network equations for primitive paths $(A,S,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma$ if for $i\in\mathcal{J}$, X_i,Z_i,Y_i and B_i are non-decreasing, Q_i is non-negative, $\lim_{t\to-\infty}Q_i(t)/t=0$ and the following set of coupled equations are satisfied:*

$$X_i = A_i+\sum_{j\in\mathcal{J}} R_{i,j}\circ Z_j, \tag{5}$$

$$Z_i = S_i\circ B_i, \tag{6}$$

$$Q_i = X_i - Z_i, \tag{7}$$

$$\int_{-\infty}^{\infty} Q_i(s)dY_i(s) = 0, \tag{8}$$

$$B_i = \iota - Y_i. \tag{9}$$

Given $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$, we let $\mathcal{Q}(A, S, R)$ be the set of quintuples (X, Z, Q, Y, B) satisfying equations (5)–(9).

The relations (5)–(9) are intuitive. For example, the first equation states that the cumulative arrivals into queue i up to time t are the sum of the customers that arrived into node i exogenously and the number of customers routed from all other queues to queue i up to time t, while the equation (8) states that a queue idles if and only if it is empty. Similar relations appear in many models of the behavior of open, single-class queueing networks [8].

The network equations presented above are more general than required to model a Jackson network as they allow the cumulative number of customers to be negative (in order to model the infinite past) and fractional (in order to permit scaling and to describe fluid limit trajectories). In particular, when the network primitive S is taken to be continuous, then served customers are counted partially in both the departures and queue lengths. This more general formulation, on the entire real line $\mathbb{R}$, was investigated in [27] and will be used in our subsequent analysis.

We say that a random process Q in $\mathcal{D}^n$ is stationary if there exists a probability measure ν on $\mathbb{R}^n_+$ such that $Q(t) \sim \nu$ for every $t \in \mathbb{R}$.

Proposition 2 *Given the Jackson network primitives $(\tilde{A}, \tilde{S}, \tilde{R})$ described above, there exist stochastic processes $\tilde{X} \in \mathcal{I}_\lambda$, $\tilde{Z} \in \mathcal{I}_\lambda$, $\tilde{Q} \in \mathcal{D}_0$, $\tilde{Y} \in \mathcal{I}_{e-\varrho}$, and $\tilde{B} \in \mathcal{I}_\varrho$ having the following three properties:*

1. $(\tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \in \mathcal{Q}(\tilde{A}, \tilde{S}, \tilde{R})$ *a.s.*

2. $\int_{\mathbb{R}} 1_{\{\tilde{Q}_i = 0\}} d\tilde{B}_i = 0$ *for every* $i \in \mathcal{J}$, *a.s.*

3. $\tilde{Q}$ *is stationary.*

Proof. The statement of this proposition is a direct consequence of Theorem 5 in [27]. ■

We will refer to the process $(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B})$ with the properties given in Proposition 2 as the Jackson network process (associated with arrival rates α, service rates μ and routing rates γ). It is not difficult to check that $\tilde{Q}$ is a positively recurrent Markov process with countable state space $\mathbb{Z}^n_+$ and discontinuous transition probabilities at the boundary [23]. It is known [23] that the distribution ν of the stationary queue length process $\tilde{Q}$ is unique and has the form

$$\nu(\kappa) = \mathbb{P}\left(\tilde{Q}(0) = \kappa\right) \quad = \quad \prod_{i \in \mathcal{J}} \frac{\mu_i - \lambda_i}{\mu_i} \left(\frac{\lambda_i}{\mu_i}\right)^{\kappa_i}.$$

for $\kappa \in \mathbb{Z}_+^n$.

Definition 1 does not directly imply that condition 2 of Proposition 2 is satisfied because although the condition that an empty queue cannot be busy is valid for discrete customer models in which the network primitives increase purely through jumps, it is not satisfied in fluid models, in which the network primitives are continuous and empty queues can be partially loaded. This subtle difference is also captured in equations (2.2)–(2.3) and conditions 2.7.A to 2.7.C in [7].

We close this section with a simple remark on the behavior of solutions to the network equations associated with continuous and absolutely continuous primitives, which we will be used in our subsequent analysis. For a proof, see Theorem 1 and Lemma 16 in [27]. The last statement follows by integration. We recall the general convention to use lower case $f = \dot{F}$ to denote the (componentwise) derivative of an upper case absolutely continuous path F.

Remark 3 Any $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\mu$ satisfies the following properties:

(a) If S is continuous, then $\emptyset \neq \mathcal{Q}(A, S, R) \subseteq \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho$.

(b) If (A, S, R) is absolutely continuous, then every $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ is absolutely continuous and its derivatives satisfy for a.e. $t \in \mathbb{R}$ and $i \in \mathcal{J}$,

$$\begin{array}{rclcrcl} x_i(t) &=& a_i(t) + \sum_{j \in \mathcal{J}} r_{i,j}\left(Z_j(t)\right) z_j(t), & & b_i(t) &=& 1 - y_i(t), \\ z_i(t) &=& s_i\left(B_i(t)\right) b_i(t), & & Q_i(t) > 0 &\Rightarrow& y_i(t) = 0, \\ q_i(t) &=& x_i(t) - z_i(t), & & Q_i(t) = 0 &\Rightarrow& q_i(t) = 0. \end{array} \tag{10}$$

(c) Lastly, suppose that A, S, R, X, Z, B, Y are absolutely continuous and nondecreasing, and Q is absolutely continuous and non-negative. If these functions satisfy equations (10) for almost every $t \in \mathbb{R}$ and $i \in \mathcal{J}$, and, in addition, for every $i \in \mathcal{J}$,

$$\begin{array}{rclrcl} X_i(0) &=& A_i(0) + \sum_{j \in \mathcal{J}} R_{i,j}(Z_j(0)), & Z_i(0) &=& S_i(B_i(0)), \\ Q_i(0) &=& X_i(0) - Z_i(0), & B_i(0) &=& -Y_i(0), \end{array} \tag{11}$$

then $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$.

3 Main Result

In Section 3.2 we present the main result of the paper, Theorem 7, which describes the asymptotically "most likely" way in which large stationary queues build up. The dynamics of these paths is characterized by a set of coupled ordinary differential equations, which is introduced in Section 3.1. An illustrative example of the minimizing large deviations trajectory in a 3-node Jackson network is presented in Section 3.3.

3.1 Description of large deviation trajectories

We now introduce some ordinary differential equations that will be shown in Theorem 7 to describe the most likely way in which the stationary queue builds up to a given level $Q^0 \in \mathbb{R}_+^n$ at time 0 and then drains.

First, for $N \subseteq \mathcal{J}$, we define the vectors $\xi(N)$ and $\zeta(N) \in \mathbb{R}_+^n$ by

$$\xi(N) := \alpha + \gamma\zeta(N), \tag{12}$$

$$\zeta_i(N) := \begin{cases} \mu_i & \text{if } i \in N, \\ \min\{\mu_i, \xi_i(N)\} & \text{otherwise.} \end{cases} \tag{13}$$

Here, $\xi_i(N)$ represents the total arrival rate into queue i when the set of non-empty queues in the system is N, while $\zeta_i(N)$ represents the total departure rate from queue i when the set of non-empty queues is N. Next, for $i, j \in \mathcal{J}$, we define

$$\bar{\alpha}_i := \gamma_{0,i}\lambda_i, \tag{14}$$

$$\bar{\gamma}_{i,j} := \gamma_{j,i}\frac{\lambda_i}{\lambda_j}, \tag{15}$$

and let the vectors $\bar{\xi}(N)$ and $\bar{\zeta}(N) \in \mathbb{R}^n$ be defined as in (12) and (13), respectively, but with α and γ replaced by $\bar{\alpha}$ and $\bar{\gamma}$, respectively. We note that the values in (14) and (15), respectively, define arrival and routing rates in the time-reversed network, which is again a Jackson network, as displayed in Figure 2 (see, for example, [24]). The matrix $\bar{\gamma}$ is substochastic and satisfies the equation

$$\lambda = \bar{\alpha} + \bar{\gamma}\lambda. \tag{16}$$

The following properties of these vectors will be needed and are taken from Lemma 16 of [26].

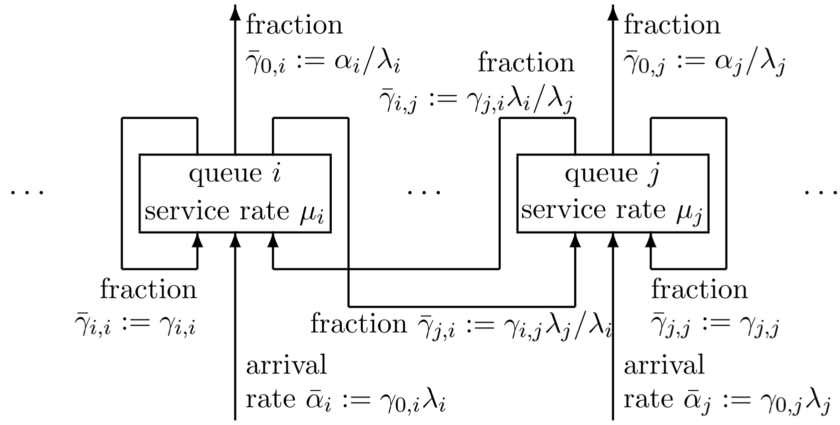

Figure 2: Time-reversed network

Lemma 4 *For every $N \subseteq \{1, \dots, n\}$, there exists a unique solution to the equations (12) and (13). The solution satisfies $\xi(N) \geq \lambda$ and $\zeta(N) \geq \lambda$, with equality holding if and only if $N = \emptyset$. For every $N \neq \emptyset$, there exists $i \in N$ such that $\xi_i(N) - \zeta_i(N) < 0$. If $M \subseteq N$ then $\zeta(M) \leq \zeta(N)$ and $\xi(M) \leq \xi(N)$. Analogous properties hold for the vectors $\bar\xi(N)$ and $\bar\zeta(N)$.*

We define the absolutely continuous queue lengths function Q^* through the relations

$$\begin{array}{rcll} Q^*(0) & = & Q^0, & \\ q^*(t) & = & \xi(N^*_t) - \zeta(N^*_t), & \text{for almost all } t > 0, \\ q^*(t) & = & \bar\zeta(N^*_t) - \bar\xi(N^*_t), & \text{for almost all } t < 0, \\ N^*_t & = & \{i \in \mathcal{J}\colon Q^*_i(t) > 0\}, & \text{for all } t \in \mathbb{R}. \end{array} \tag{17}$$

These coupled differential equations possess a piecewise linear solution which can be calculated via the following algorithm:

1. Initialization: set $t_0 := 0$, $N_0 := \mathcal{J}$ and $i = 1$; define $N^*_0 := \{k \in \mathcal{J}\colon Q^0_k > 0\}$ and $Q^*(0) := Q^0$.

2. Define the vectors $\xi^i := \xi(N_{i-1})$, $\zeta^i := \zeta(N_{i-1})$, $\xi^{-i} := \bar\xi(N_{-i+1})$, $\zeta^{-i} := \bar\zeta(N_{-i+1})$ by solving equations (12) and (13).

3. If there exists $j \in \mathcal{J} \setminus N_{i-1}$ with $Q^i_j = 0$ and $\xi^i_j - \zeta^i_j \leq 0$, set $t_i := t_{i-1}$. Otherwise, define t_i to be the solution of the linear minimization problem

$$t_i \;:=\; \sup\left\{t \geq t_{i-1} : Q(t_{i-1}) + (\xi^i - \zeta^i)(t - t_{i-1}) \geq 0\right\}.$$

Similarly, if there exists $j \in \mathcal{J} \setminus N_{-i+1}$ with $Q_j^{-i} = 0$ and $\xi_j^{-i} - \zeta_j^{-i} \leq 0$, set $t_{-i} := t_{-i+1}$. Otherwise, define t_{-i} to be the solution of the linear minimization problem

$$t_{-i} \;:=\; \inf\left\{t \leq t_{-i+1} : Q(t_{-i+1}) + (\xi^{-i} - \zeta^{-i})(t_{-i+1} - t) \geq 0\right\}.$$

4. If $t_i > t_{i-1}$, define $N_t^* := N_i$ and $Q^*(t) := Q(t_{i-1}) + (\xi^i - \zeta^i)(t - t_{i-1})$ for $t \in [t_{i-1}, t_i)$, $t \neq 0$.

   Likewise, if $t_{-i} < t_{-i+1}$, define $N_t^* := N_{-i}$ and $Q^*(t) := Q(t_{-i+1}) + (\xi^{-i} - \zeta^{-i})(t_{-i+1} - t)$ for $t \in (t_{-i}, t_{-i+1}]$, $t \neq 0$.

5. Set $Q^i := Q^{i-1} + (\xi^i - \zeta^i)(t_i - t_{i-1})$ and $Q^{-i} := Q^{-i+1} + (\xi^{-i} - \zeta^{-i})(t_{-i} - t_{-i+1})$.

6. Choose $j \in N_{i-1}$ such that $Q_j^i = 0$ and $\xi_j^i \leq \zeta_j^i$, and define $N_i := N_{i-1} \setminus \{j\}$.

   Choose $k \in N_{-i+1}$ such that $Q_k^{-i} = 0$ and $\xi_k^{-i} \leq \zeta_k^{-i}$, and define $N_{-i} := N_{-i+1} \setminus \{k\}$.

7. Termination: if $i < n$, increment i by one and repeat the algorithm, starting from step 2. Otherwise, define $N_t^* := \emptyset$ and $Q^*(t) := 0$ for $t \in (-\infty, t_{-n}] \cup [t_n, \infty)$, and terminate the algorithm.

In view of Lemma 4, the minimization problems in step 3 have unique finite solutions, and step 3 ensures that j and k in step 6 exist. Step 4 ensures that the middle two equations of display (17) are satisfied. In addition, if $j \in \mathcal{J} \setminus N_i$ for $i \in \{-n, -n+1, \ldots, n\}$ then $Q_j^i = 0$ and $\xi_j^i = \zeta_j^i$. In particular, the queue length vectors $Q^*(t)$ and sets N_t^* are well defined for every $t \in \mathbb{R}$ and satisfy the conditions of display (17). Furthermore, our next lemma asserts that this solution is unique.

Lemma 5 *The solution $Q^* \in \mathcal{D}_0$ to the set of equations (17) is unique. In addition, Q^* is piecewise linear, non-negative and satisfies $Q^*(t) = 0$ whenever $|t|$ is sufficiently large.*

Proof. Suppose Q^* is a solution to the set of equations (17). Along with the definition of ξ and ζ given in (12) and (13), these conditions imply for almost every $t \in \mathbb{R}_+$,

$$Q_i^*(-t) < 0 \Rightarrow i \in \mathcal{J} \setminus N_{-t}^* \Rightarrow \bar{\zeta}_i(N_{-t}^*) \leq \bar{\xi}_i(N_{-t}^*) \Rightarrow q_i^*(-t) \leq 0.$$

Hence, the inequality $Q_i^*(-t) < 0$ contradicts the property that $Q_i^*(0) = Q_i^0 \geq 0$. We conclude that $Q^*(-t) \geq 0$ for every $t \in \mathbb{R}_+$.

For $t \geq 0$, we define $X(t) := Q^0 + \bar{\alpha}t - (I - \bar{\gamma})\mathrm{diag}(\mu)t$, $Q(t) := Q^*(-t)$ and $Y(t) := (I - \bar{\gamma})^{-1}(Q(t) - X(t))$ in $\mathbb{R}^n$. Clearly, $Q(t) = X(t) + (I - \bar{\gamma})Y(t) = Q^*(-t) \geq 0$ for $t \geq 0$. Furthermore, Y is absolutely continuous and, by (17), its derivatives satisfy

$$y(t) \;=\; (I - \bar{\gamma})^{-1}\left(\bar{\xi}(N_t) - \bar{\zeta}(N_t) - \bar{\alpha} + (I - \bar{\gamma})\mu\right) = \mu - \bar{\zeta}(N_t)$$

for almost all $t \geq 0$, with $N_t := \{i \in \mathcal{J} \colon Q_i(t) > 0\}$. Since $\zeta_i(N_t) = \mu_i$ if $Q_i(t) > 0$, this implies for every $i \in \mathcal{J}$

$$\int_{\mathbb{R}_+} Q_i(t) dY_i(t) \;=\; \int_{\mathbb{R}_+} Q_i(t) y_i(t)\, dt = 0.$$

Hence the pair (Q, Y) is a solution to the Skorokhod problem for X with reflection matrix $I - \bar{\gamma}$ [21]. Since $\bar{\gamma}$ has a spectral radius strictly less than 1, this problem is known to have a unique solution [21] (see also [15]). In particular, the values $Q^*(t), t \leq 0$, attained by any solution Q^* to the equations (17), are uniquely determined. In a similar fashion, one can also prove the uniqueness of Q^* on $\mathbb{R}_+$ and, thus, on all of $\mathbb{R}$. ■

Given the unique solution Q^* to the set of equations (17), and the associated processes N^*, $\xi(N^*)$ and $\zeta(N^*)$, we define the absolutely continuous functions A^*, S^*, R^*, X^*, Z^*, Y^*, B^* by setting their values at 0 to be

$$\begin{array}{rclrcl} A^*(0) & = & 0, & S^*(0) & = & 0, \\ R^*(0) & = & 0, & X^*(0) & = & -\gamma(I-\gamma)^{-1}Q^0, \\ Z^*(0) & = & -(I-\gamma)^{-1}Q^0, & Y_i^*(0) & = & -\mathrm{diag}(\mu)^{-1}Z^*(0), \\ B_i^*(0) & = & \mathrm{diag}(\mu)^{-1}Z^*(0); & & & \end{array} \tag{18}$$

specifying the derivatives for a.e. $t > 0$ and $i \in \mathcal{J}$, to be

$$\begin{array}{rclrcl} a_i^*(t) & = & \alpha_i, & s_i^*(B_i^*(t)) & = & \mu_i, \\ r_{i,j}^*(Z_j^*(t)) & = & \gamma_{i,j}, & x_i^*(t) & = & \xi_i(N_t^*), \\ z_i^*(t) & = & \zeta_i(N_t^*), & y_i^*(t) & = & 1 - \zeta_i(N_t^*)/\mu_i, \\ b_i^*(t) & = & \zeta_i(N_t^*)/\mu_i; & & & \end{array} \tag{19}$$

and defining the derivatives, for a.e. $t < 0$ and $i \in \mathcal{J}$, to be

$$\begin{array}{rclrcl} a_i^*(t) & = & \bar{\alpha}_i\bar{\zeta}_i(N_t^*)/\lambda_i, & s_i^*(B_i^*(t)) & = & \mu_i\bar{\xi}_i(N_t^*)/\bar{\zeta}_i(N_t^*), \\ r_{i,j}^*(Z_j^*(t)) & = & \bar{\gamma}_{j,i}\bar{\zeta}_i(N_t^*)/\bar{\xi}_j(N_t^*), & x_i^*(t) & = & \bar{\zeta}_i(N_t^*), \\ z_i^*(t) & = & \bar{\xi}_i(N_t^*), & b_i^*(t) & = & \bar{\zeta}_i(N_t^*)/\mu_i, \\ y_i^*(t) & = & 1 - \bar{\zeta}_i(N_t^*)/\mu_i. & & & \end{array} \tag{20}$$

By Lemma 4 and the relation (3), we know that for every $i \in \mathcal{J}$ and $t \in \mathbb{R}$,

$$\xi_i(\emptyset) = \zeta_i(\emptyset) = \lambda_i \leq \xi_i(N_t^*) \wedge \zeta_i(N_t^*). \tag{21}$$

Since $\lambda > 0$, this implies $b^*(t) > 0$ and $z^*(t) > 0$ for every $t \in \mathbb{R}$, and so the equations (18)–(20) have a unique solution. The intuition behind the form of these differential equations is provided in Section 3.2 and an illustrative example is presented in Section 3.3. We close this section by showing that $(X^*, Z^*, Q^*, Y^*, B^*)$ is the unique solution to the network equations associated with (A^*, S^*, R^*).

Lemma 6 $(A^*, S^*, R^*) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$ *and* $\mathcal{Q}(A^*, S^*, R^*) = \{(X^*, Z^*, Q^*, Y^*, B^*)\}$.

Proof. In view of Lemma 5, $N_t^* = \emptyset$ for all $|t|$ sufficiently large. From the equalities in (21) and the relations for b^* and z^* specified in (19) and (20), if follows that $B^* \in \mathcal{I}_\varrho$, $Z^* \in \mathcal{I}_\lambda$ and, since $\lambda > 0$ and $\varrho > 0$, that the ranges of B^* and Z^* are both equal to $\mathbb{R}$. Together, the last two statements ensure that S^* and R^* are well-defined on all of $\mathbb{R}$ and that $S^* \in \mathcal{I}_\mu \cap \mathcal{C}_\mu$, $A^* \in \mathcal{I}_\alpha$ and $R^* \in \mathcal{I}_\gamma$. This shows that $(A^*, S^*, R^*) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$.

We now show that $(X^*, Z^*, Q^*, Y^*, B^*) \in \mathcal{Q}(A^*, S^*, R^*)$. Since (A^*, S^*, R^*) is absolutely continuous by definition, by Remark 3(c), it suffices to show that (10) and (11) are satisfied, with all quantities replaced by their starred versions. By (18), $A^*(0) = 0$, $R^*(0) = 0$ and $Z^*(0) = -(1-\gamma)^{-1}Q^0 \leq 0$, where the latter inequality holds because $Q^0 = Q^*(0) \geq 0$ by Lemma 5 and $(I - \gamma)^{-1}$ is non-negative, and by (19) and the fact that Z_j^* is non-decreasing, we have $r_{i,j}^*(t) = \gamma_{i,j}$ for a.e. $t \in [Z_j^*(0), 0]$. Together, the last two statements imply $R_{i,j}(Z_j^*(0)) = \gamma_{i,j} Z_j^*(0)$ for every $i, j \in \mathcal{J}$, which when combined with the definition of $X^*(0)$ in (18), shows that $X^*(0) = \gamma Z^*(0)$ and $Q^*(0) = Q^0 = X^*(0) - Z^*(0)$. This immediately implies the two relations on the left-hand side of (11). In a similar fashion, we also have $B^*(0) = diag(\mu)^{-1} Z^*(0) = -Y^*(0) \leq 0$ and $S^*(0) = 0$ by (18) and $s_i^*(t) = \mu_i$ for a.e. $t \in [B_i^*(0), 0]$ by (19). This implies $S_i^*(B_i^*(0) = \mu_i B_i^*(0) = Z_i^*(0)$ for every $i \in \mathcal{J}$, thus completing the proof of (11).

We now turn to establishing (10). The second, third and fifth relations in (10) are an immediate consequence of equations (17), (19) and (20). The first relation in (10) holds because, from the definition (12) of $\xi(N^*)$ and the definitions of $a_i^*, r_{i,j}^*, Z_j^*$ and z_j^* given in (19), we have for $t \geq 0$

$$x_i^*(t) = \xi_i(N_t^*) = \alpha_i + \sum_{i \in \mathcal{J}} \gamma_{i,j} \zeta_j(N_t^*) = a_i^*(t) + \sum_{i \in \mathcal{J}} r_{i,j}^*(Z_j^*(t)) z_j^*(t),$$

and, by (2), (14) and the definitions of $a_i^*, r_{i,j}^*, Z_j^*$ and z_j^* given in (20), we have for $t \le 0$,

$$\begin{aligned} x_i^*(t) = \bar\zeta_i(N_t^*) &= \left(\gamma_{0,i} + \sum_{j\in\mathcal{J}} \gamma_{j,i}\right)\bar\zeta_i(N_t^*) = \frac{\bar\alpha_i}{\lambda_i}\bar\zeta_i(N_t^*) + \sum_{j\in\mathcal{J}} \gamma_{j,i}\bar\zeta_i(N_t^*) \\ &= a_i^*(t) + \sum_{j\in\mathcal{J}} r_{i,j}^*(Z_j^*(t))z_j^*(t). \end{aligned}$$

Next, for $t \in \mathbb{R}$ such that $Q_i^*(t) > 0$, we have $i \in N_t^*$ and hence, by (13), $\zeta_i(N_t^*) = \mu_i$ if $t \ge 0$ and $\bar\zeta_i(N_t^*) = \mu_i$ if $t \le 0$. From the definition of y_i^* given in (19) and (20), it follows that $y_i^*(t) = 0$, and so the fourth relation in (10) also follows, concluding the proof that $(X^*, Z^*, Q^*, Y^*, B^*) \in \mathcal{Q}(A^*, S^*, R^*)$.

Since the components of (A^*, S^*, R^*) are piecewise linear, the statement that $(X^*, Z^*, Q^*, Y^*, B^*)$ is the unique element of $\mathcal{Q}(A^*, S^*, R^*)$ follows from Lemma 19 of [26]. The prerequisites therein follow from equations (3) and (16). This completes the proof. ■

3.2 Statement of the Main Result

Recall the primitives and network processes $(\tilde A, \tilde S, \tilde R, \tilde X, \tilde Z, \tilde Q, \tilde Y, \tilde B)$ associated with the Jackson network, as described in Section 2. From Proposition 2, it immediately follows that $(\tilde A, \tilde S, \tilde R, \tilde X, \tilde Z, \tilde Q, \tilde Y, \tilde B)$ lies in the space

$$\mathcal{M} := \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma \times \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho.$$

Let Γ_k be the 'usual "functional law of large numbers" scaling transformation of functions i.e., we have

$$(\Gamma_k D)(t) \quad = \quad \frac{D(kt)}{k} \qquad \text{for } t \in \mathbb{R}.$$

We are now in a position to state the main result of this work.

Theorem 7 *The sequence of scaled processes $\Gamma_k(\tilde A, \tilde S, \tilde R, \tilde X, \tilde Z, \tilde Q, \tilde Y, \tilde B)$, conditioned on $\tilde Q(0) \ge kQ^0$, converges in distribution in $\mathcal{M}$ to the Dirac measure of the path $(A^*, S^*, R^*, X^*, Z^*, Q^*, Y^*, B^*)$, as $k \to \infty$. This convergence is exponentially fast in the sense that for every $\epsilon > 0$,*

$$\begin{aligned} \limsup_{k\to\infty} \frac{1}{k} \log P\Big(\Big\| &\Gamma_k(\tilde A, \tilde S, \tilde R, \tilde X, \tilde Z, \tilde Q, \tilde Y, \tilde B) \\ &- (S^*, A^*, R^*, X^*, Y^*, B^*, Z^*, Q^*)\| \ge \epsilon \mid \tilde Q(0) \ge kQ^0\Big) \quad < \quad 0. \quad (22) \end{aligned}$$

The proof of the theorem is presented in Section 5.5. In contrast to the results in [2], our proof is carried out entirely within the framework of large deviations theory. Also, the exponentially fast convergence stated in Theorem 7 is stronger than the probabilistic bounds given in [2].

We note that the differential equations in Section 3.1, which describe the asymptotic behavior on the positive time interval, simply characterize the way in which a fluid Jackson network drains when it starts at level Q^0 at time 0, and arrival, service and routing rates are given by the mean rates α, μ, and γ, respectively. This asymptotic conditional behavior is well known from functional laws of large numbers and is referred to as the fluid limit (see [6]). The important part in the theorem is the description of the asymptotic conditional behavior on the negative time interval. Here, the differential equations describe the way in which queues build up, and as we will show in Section 4, this is equivalent to the way in which a time-reversed queueing network with arrival rates $\bar{\alpha}$, service rates μ and routing rates $\bar{\gamma}$ drains from a level Q_0, but backwards in time. In this backwards description, the roles of the departures from and entries to the network are interchanged.

3.3 A 3-dimensional Example

In order to illustrate the explicit nature of our results, we consider a Jackson network with $n = 3$ queues, external arrival rate vector $\alpha = (2, 0, 0)^T$, service rate vector $\mu = (3, 3, 5)^T$ and routing matrix

$$\gamma = \begin{pmatrix} 0 & 0 & 0.7 \\ 0.8 & 0 & 0.2 \\ 0 & 0.4 & 0 \end{pmatrix}.$$

This network is depicted in Figure 3.

We calculated the paths A^*, S^*, R^*, X^*, Z^*, Q^* Y^*, B^*, for the target queue lengths $Q^0 := (0, 0, 1)^T$ via the algorithm specified in Section 3.1. In Figure 4, we show the components of Q^*. Observe that three changes of slope are required to build up the queue lengths Q^0 at time zero and another three changes of slopes are required to drain them. In Figure 5, we plot the difference between the cumulative arrivals A_1^* and the average cumulative arrivals $\alpha_1 \iota$, and the difference between the cumulative number of served customers $S_1^* \circ B_1^*$, $S_2^* \circ B_2^*$, $S_3^* \circ B_3^*$ and the corresponding average cumulative number of served customers $\lambda_1 \iota$, $\lambda_2 \iota$, $\lambda_3 \iota$. Finally, in Figure 6, we

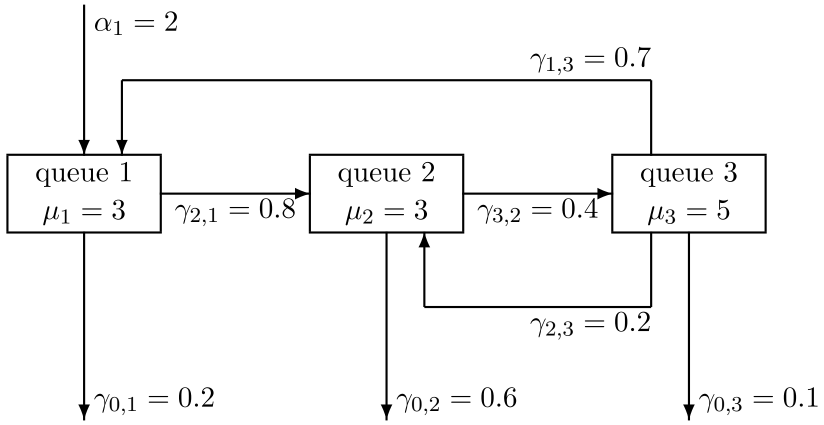

Figure 3: Example network

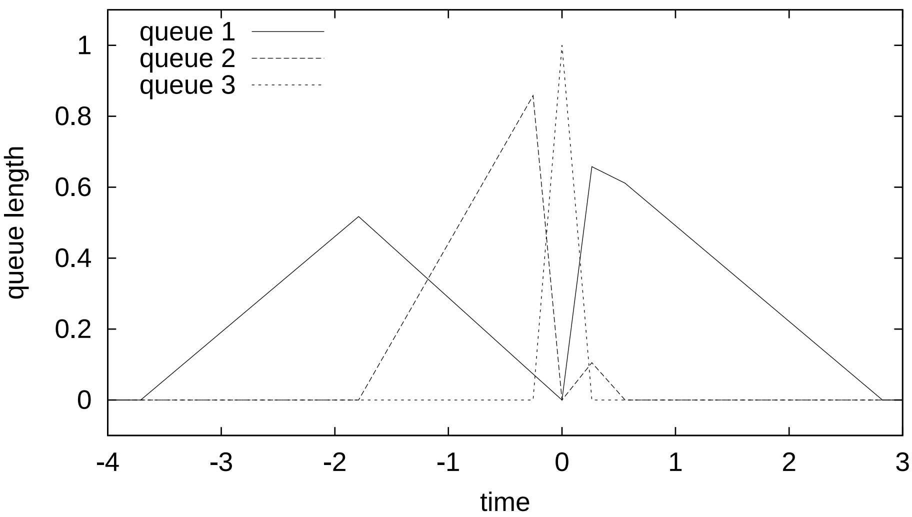

Figure 4: Queue lengths Q^*

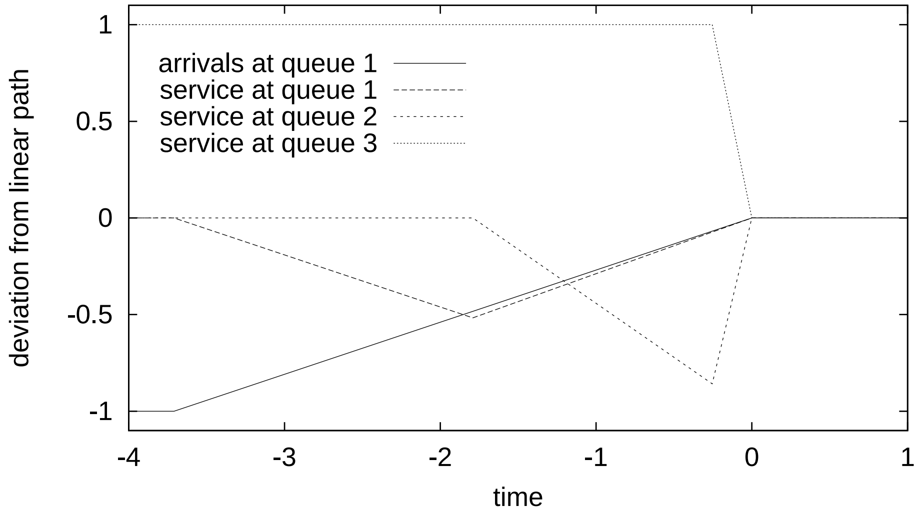

Figure 5: Difference of cumulative external arrivals and number of served customers from averages

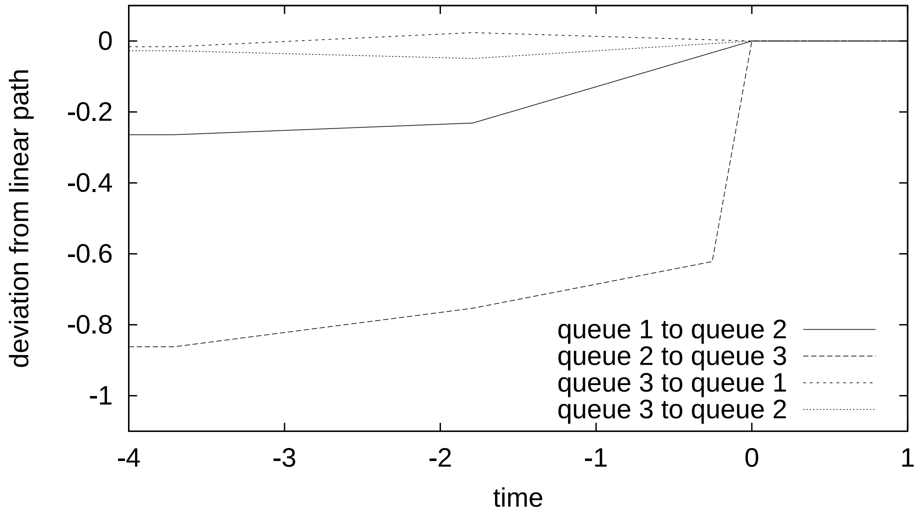

Figure 6: Difference of cumulative routing decisions from averages

illustrate the difference between the cumulative routing decisions $R^*_{1,2} \circ S^*_2 \circ B^*_2$, $R^*_{2,3} \circ S^*_3 \circ B^*_3$, $R^*_{3,1} \circ S^*_1 \circ B^*_1$, $R^*_{3,2} \circ S^*_2 \circ B^*_2$ and the corresponding average cumulative routing decisions $\gamma_{1,2}\lambda_2\iota$, $\gamma_{2,3}\lambda_3\iota$, $\gamma_{3,1}\lambda_1\iota$, $\gamma_{3,2}\lambda_2\iota$. We note that after time 0, these differences are zero.

This example shows that a coincidence of a complicated pattern of deviations from the mean behavior of the arrival, service and routing processes characterizes the most likely way in which a large queue length at queue 3 builds up. Our results permit the explicit calculation of such patterns in general Jackson networks.

4 Time-reversed network behavior

In (14)–(15) we introduced the parameters associated with the time-reversed Jackson network. In this section, we define a more detailed, deterministic time-reversal of the Jackson network at the pathwise, functional level. Specifically, we show that for certain network primitive paths (A, S, R) and corresponding network paths $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$, we can associate corresponding time-reversed paths $(\bar{A}, \bar{S}, \bar{R})$ and $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B}) \in \mathcal{Q}(\bar{A}, \bar{S}, \bar{R})$. The time-reversed network will be used to identify the most likely large deviation trajectories in Section 5.4. We start by defining a class of primitive paths and associated network paths that are amenable to time-reversal.

Definition 8 (Reversibility) *Any pair of network primitive paths $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$ and associated network paths $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ are said to be reversible if and only if they satisfy the following three properties:*

(a) A, S and R are continuous;

(b) for every $i \in \mathcal{J}$, the path

$$R_{0,i} \;:=\; \iota - \sum_{j \in \mathcal{J}} R_{j,i} \tag{23}$$

is non-decreasing;

(c) there exists a vector $\eta \in \mathbb{R}^n$ satisfying

$$X_i(\eta_i) \;=\; R_{0,i}(Z_i(0)) + \sum_{j \in \mathcal{J}} R_{j,i}(Z_i(\eta_j)) \quad \text{for } i \in \mathcal{J}. \tag{24}$$

Given reversible network primitives $(A,S,R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$, associated network processes $(X,Z,Q,Y,B) \in \mathcal{Q}(A,S,R)$, and a vector η satisfying condition (24), we now define the corresponding time-reversed primitive paths $(\bar{A},\bar{S},\bar{R})$ and associated network paths $(\bar{X},\bar{Z},\bar{Q},\bar{Y},\bar{B})$. The first property that must be satisfied by the time-reversed paths, which justifies the nomenclature, is

$$\bar{Q}(t) \quad := \quad Q(-t) \tag{25}$$

for every $t \in \mathbb{R}$. In addition, for $i,j \in \mathcal{J}$ and $t \in \mathbb{R}$, we define

$$\begin{aligned}
\bar{X}_i(t) &:= X_i(\eta_i) - Z_i(-t), && (26)\\
\bar{Z}_i(t) &:= X_i(\eta_i) - X_i(-t), && (27)\\
\bar{R}_{i,j}(\bar{Z}_j(t)) &:= R_{j,i}(Z_i(\eta_j)) - R_{j,i}(Z_i(-t)), && (28)\\
\bar{A}_i(t) &:= Z_i(0) - Z_i(-t) - \sum_{j\in\mathcal{J}} R_{j,i}(Z_i(0)) + \sum_{j\in\mathcal{J}} R_{j,i}(Z_i(-t)), && (29)\\
\bar{B}_i(t) &:= B_i(\eta_i) - B_i(-t), && (30)\\
\bar{Y}_i(t) &:= t - \bar{B}_i(t), && (31)\\
\bar{S}_i(\bar{B}_i(t)) &:= \bar{Z}_i(t). && (32)
\end{aligned}$$

The following theorem establishes certain basic properties of the time-reversed paths.

Theorem 9 *Given reversible network primitive paths $(A,S,R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$ and network paths $(X,Z,Q,Y,B) \in \mathcal{Q}(A,S,R)$, the time-reversed paths $(\bar{A},\bar{S},\bar{R})$ and $(\bar{X},\bar{Z},\bar{Q},\bar{Y},\bar{B})$ are well-defined by equations (25)–(32). Moreover, $(\bar{A}(0),\bar{S}(0),\bar{R}(0)) = (0,0,0)$, $(\bar{A},\bar{S},\bar{R}) \in \mathcal{I}_{\bar{\alpha}} \times \mathcal{I}_\mu \times \mathcal{I}_{\bar{\gamma}}$ and $(\bar{X},\bar{Z},\bar{Q},\bar{Y},\bar{B}) \in \mathcal{Q}(\bar{A},\bar{S},\bar{R}) \cap \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho$.*

Proof. Since Definition 8(a) ensures that (A,S,R) is continuous, from Remark 3(a), it follows that $(X,Z,Q,Y,B) \in \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho$. From the definitions (25)–(27) and (29)–(31), it is clear that the processes $\bar{X},\bar{Z},\bar{Q},\bar{Y},\bar{B}$ and $\bar{A}$ are well-defined and, moreover, that $\bar{X},\bar{Z} \in \mathcal{I}_\lambda$, $\bar{Q} \in \mathcal{D}_0$, $\bar{Y} \in \mathcal{I}_{e-\varrho}$ and $\bar{B} \in \mathcal{I}_\varrho$. In addition, the equations (29) and (23) imply that for $i \in \mathcal{J}$,

$$\bar{A}_i(t) = R_{0,i}(Z_i(0)) - R_{0,i}(Z_i(-t)).$$

In turn, this implies that $\bar{A}(0) = 0$ and, since $R_{0,i} \in \mathcal{I}_{\gamma_{0,i}}$ and $Z_i \in \mathcal{I}_{\lambda_i}$ for every $i \in \{1,\dots,n\}$, that $\bar{A} \in \mathcal{I}_{\bar{\alpha}}$.

We shall now argue that $\bar{R}$ is well-defined by equation (28), satisfies $\bar{R}(0) = 0$ and lies in $\mathcal{I}_{\bar{\gamma}}$. Fix $i, j \in \mathcal{J}$. Since $\bar{Z}_j$ is non-decreasing, it suffices to show that for any $s < t$ such that $\bar{Z}_j(t) = \bar{Z}_j(s)$, we have

$$R_{j,i}(Z_i(-s)) = R_{j,i}(Z_i(-t)).$$

By (26) and (27), $\bar{Z}_j(t) = \bar{Z}_j(s)$ implies that $X_j(-s) = X_j(-t)$ and thus the equality in the last display follows from the network equation (5) and the fact that A_j, $R_{j,k}$ and Z_k, $k \in \mathcal{J}$, are all non-decreasing. Thus $\bar{R}_{i,j}$ is well-defined. Moreover, since $\bar{Z}_j(-\eta_j) = 0$ by (27), we have

$$\bar{R}_{i,j}(0) \;=\; \bar{R}_{i,j}(\bar{Z}_j(-\eta_j)) = R_{j,i}(Z_i(\eta_j)) - R_{j,i}(Z_i(\eta_j)) = 0.$$

Since $R_{j,i} \in \mathcal{I}_{\gamma_{j,i}}$ and $Z_i \in \mathcal{I}_\lambda$, it follows from (28) and the definition of $\bar{\gamma}$ given in (15) that $\bar{R}$ is an element of $\mathcal{I}_{\bar{\gamma}}$.

We now turn to the definition of $\bar{S}$. Fix $i \in \mathcal{J}$, and let $s < t$ be such that $\bar{B}_i(s) = \bar{B}_i(t)$. Then equations (30) and (6) imply $B_i(-t) = B_i(-s)$ and $Z_i(-t) = Z_i(-s)$. By (9), this implies Y_i is strictly increasing on $[-t, -s]$ which, along with the “complementarity condition” (8), implies that $Q_i(-t) = Q_i(-s) = 0$. Therefore, by (5), $X_i(-t) = Z_i(-t) = Z_i(-s) = X_i(-s)$, and so equation (27) shows that $\bar{Z}_i(t) = \bar{Z}_i(s)$. Thus $\bar{S}_i$ is well-defined by (32). Furthermore, (30), (32) and (27) together imply that $\bar{B}_i(-\eta_i) = 0$ and so $\bar{S}_i(0) = \bar{S}_i(\bar{B}_i(-\eta_i)) = \bar{Z}_i(-\eta_i) = 0$. When combined with the fact that $\bar{Z} \in \mathcal{I}_\lambda$, $\bar{B} \in \mathcal{I}_\varrho$ and the definition (4) of ϱ, it follows that $\bar{S} \in \mathcal{I}_\mu$. Thus we have established the first statement of the theorem and also shown that $(\bar{A}(0), \bar{S}(0), \bar{R}(0)) = (0, 0, 0)$ and $(\bar{A}, \bar{S}, \bar{R}) \in \mathcal{I}_{\bar{\alpha}} \times \mathcal{I}_\mu \times \mathcal{I}_{\bar{\gamma}}$ and $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B}) \in \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho$.

It only remains to show that $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B}) \in \mathcal{Q}(\bar{A}, \bar{S}, \bar{R})$. First, note that, combining equations (26), (24), (23), (29) and (28), we obtain for $i \in \mathcal{J}$ and $t \in \mathbb{R}$,

$$\begin{aligned}
\bar{X}_i(t) \;&=\; Z_i(0) - Z_i(-t) - \sum_{j \in \mathcal{J}} R_{j,i}(Z_i(0)) + \sum_{j \in \mathcal{J}} R_{j,i}(Z_i(\eta_j)) \\
&=\; \bar{A}_i(t) + \sum_{j \in \mathcal{J}} \bar{R}_{i,j}(\bar{Z}_j(t)).
\end{aligned}$$

Moreover, as a result of equations (25)–(27) and (7), we have for $i \in \mathcal{J}$ and $t \in \mathbb{R}$,

$$\bar{Q}_i(t) = Q_i(-t) = X_i(-t) - Z_i(-t) = \bar{X}_i(t) - \bar{Z}_i(t).$$

Since (31), (30) and (9) imply

$$\bar{Y}_i(t) \;=\; t - B_i(\eta_i) + B_i(-t) = Y_i(\eta_i) - \eta_i - Y_i(-t),$$

we obtain

$$\int_{-\infty}^{\infty} \bar{Q}_i(s) d\bar{Y}_i(s) \;=\; \int_{-\infty}^{\infty} Q_i(s) dY_i(s) = 0,$$

where the last equality follows from (8). Comparing the last four equations and equations (31) and (32) with the network equations of Definition 1, we see that $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B}) \in \mathcal{Q}(\bar{A}, \bar{S}, \bar{R})$ and the proof of the theorem is complete. ■

We now present a characterization of the derivatives of the time-reversed quantities in the case when they are absolutely continuous, which is analogous to the characterization given for the forward system in Remark 3(b).

Remark 10 Suppose the reversible network primitive paths (A, S, R) and network paths (X, Z, Q, Y, B) are absolutely continuous. Then it follows immediately from the definitions that the constructed time-reversed paths $(\bar{A}, \bar{S}, \bar{R})$ and $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B})$ are also absolutely continuous and their derivatives (represented by the lower case letters, as usual) satisfy for almost every $t \in \mathbb{R}$ and every $i, j \in \mathcal{J}$,

$$\begin{array}{rclrcl} \bar{a}_i(-t) & = & r_{0,i}(Z_i(t)) z_i(t), & \bar{z}_i(-t) & = & x_i(t), \\ \bar{s}_i(\bar{B}_i(-t)) \bar{b}_i(-t) & = & x_i(t), & \bar{q}_i(-t) & = & -q_i(t), \\ \bar{r}_{i,j}(\bar{Z}_j(-t)) \bar{z}_j(-t) & = & r_{j,i}(Z_i(t)) z_i(t), & \bar{y}_i(-t) & = & y_i(t), \\ \bar{x}_i(-t) & = & z_i(t), & \bar{b}_i(-t) & = & b_i(t). \end{array} \tag{33}$$

Next we describe a standard situation in which paths are reversible.

Lemma 11 *Suppose $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$ is continuous, $R_{0,i}$ is nondecreasing for every $i \in \mathcal{J}$ and $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ satisfies $Q(t) = 0$ for some $t \leq 0$. Then the pair (A, S, R) and (X, Z, Q, Y, B) is reversible.*

Proof. Clearly, we only have to show the existence of a vector η satisfying condition (24) of Definition 8. We shall obtain η as the limit of a sequence $(\eta^k)_{k \in \mathbb{N}}$ that is constructed using an iterative procedure. We start with the vector $\eta^1 := 0$ and define, for $k \in \mathbb{N}$ and $i \in \mathcal{J}$, the value η_i^{k+1} by

$$\eta_i^{k+1} \;:=\; X_i^{-1}\left(Z_i(0) - \sum_{j \in \mathcal{J}} R_{j,i}(Z_i(0)) + \sum_{j \in \mathcal{J}} R_{j,i}(Z_i(\eta_j^k)) \right), \tag{34}$$

where X_i^{-1} denotes the left-continuous pseudo inverse of X_i defined by $X_i^{-1}(c) := \inf\{t \in \mathbb{R}\colon X(t) \geq c\}$.

We first note that, for every $i \in \mathcal{J}$, since X_i^{-1} is non-decreasing and $X_i \geq Z_i$ by (7), we have

$$\eta_i^2 \;=\; X_i^{-1}(Z_i(0)) \leq X_i^{-1}(X_i(0)) = \inf\{t \in \mathbb{R}\colon X(t) \geq X(0)\} \leq 0 = \eta_i^1.$$

Since Z and R are componentwise non-decreasing, definition (34) therefore implies by induction that the sequence $(\eta^k)_{k\in\mathbb{N}}$ is componentwise non-increasing.

Next, we show that this sequence is bounded below. Let $t \leq 0$ be such that $Q(t) = 0$, which exists by assumption. Then let θ be the vector defined by $\theta_i := X_i^{-1}(X_i(t))$ for $i \in \mathcal{J}$. For $i \in \mathcal{J}$, since $X_i(t) \leq X_i(0)$, we have the inequality

$$\theta_i \;\leq\; X_i^{-1}(X_i(0)) \leq 0 = \eta_i^1.$$

This implies that $\theta \leq \eta^k$ for $k = 1$. If $\eta^k \geq \theta$ for some $k \in \mathbb{N}$, then

$$\begin{aligned} X_i(\eta_i^{k+1}) &\geq R_{0,i}(Z_i(0)) + \sum_{j\in\mathcal{J}} R_{j,i}(Z_i(t)) \\ &= R_{0,i}(Z_i(0)) - R_{0,i}(Z_i(t)) + X_i(t) \geq X_i(t) \end{aligned}$$

for every $i \in \mathcal{J}$ and hence $\eta^{k+1} \geq \theta$. (Here we used the fact that $R_{0,i} \circ Z_i$ is non-decreasing.) By induction, we conclude that $\eta^k \geq \theta$ for every $k \in \mathbb{N}$. Hence the sequence $(\eta^k)_{k\in\mathbb{N}}$ converges to a vector $\eta \in \mathbb{R}^n$. Applying the function X_i to both sides of equation (34), using the convergence of the sequence $(\eta^k)_{k\in\mathbb{N}}$, the identity $X_i \circ X_i^{-1} = \iota$ and the continuity of Z and R, we conclude that η satisfies condition (24) for every $i \in \mathcal{J}$. ∎

The last result of this section connects the pathwise time-reversal described above with the parameters $\bar{\alpha}$ and $\bar{\gamma}$ of the time-reversed Jackson network introduced in (14) and (15).

Lemma 12 *The network primitives (A^*, S^*, R^*) and network behavior $(X^*, Z^*, Q^*, Y^*, B^*)$ are reversible. If $(\bar{A}^*, \bar{S}^*, \bar{R}^*)$ and $(\bar{X}^*, \bar{Z}^*, \bar{Q}^*, \bar{Y}^*, \bar{B}^*)$ denotes the time-reversed primitives and network behavior, then for almost every $t \geq 0$ and every $i \in \mathcal{J}$,*

$$\begin{aligned} \bar{a}_i^*(t) &= \bar{\alpha}_i, \\ \bar{s}_i^*(\bar{B}^*(t)) &= \mu_i, \\ \bar{r}_i^*(\bar{Z}^*(t)) &= \bar{\gamma}_i. \end{aligned}$$

Proof. We first show that for every $i \in \mathcal{J}$, the function $R^*_{0,i} := \iota - \sum_{j\in\mathcal{J}} R^*_{j,i}$ is non-decreasing. The function $R^*_{0,i}$ is absolutely continuous and for almost all $t < 0$, definitions (23), (20) and (12) show that its derivative at $Z^*_i(t)$ satisfies

$$r^*_{0,i}(Z^*_i(t)) = 1 - \sum_{j\in\mathcal{J}} r^*_{j,i}(Z^*_i(t)) = 1 - \sum_{j\in\mathcal{J}} \frac{\bar{\gamma}_{i,j}\bar{\zeta}_j(N^*_t)}{\bar{\xi}_i(N^*_t)} = 1 - \frac{\bar{\xi}_i(N^*_i) - \bar{\alpha}_i}{\bar{\xi}_i(N^*_i)} \geq 0.$$

Similarly, using (23) and (19), we have for almost all $t \geq 0$,

$$r^*_{0,i}(Z^*_i(t)) = 1 - \sum_{j\in\mathcal{J}} r^*_{j,i}(Z^*_i(t)) = 1 - \sum_{j\in\mathcal{J}} \gamma_{i,j} \geq 0,$$

where the latter follows from the assumption that γ is substochastic. This shows that $r^*_{0,i} \circ Z^*_i$ is non-decreasing. Since the equation for z^*_i in (20) shows that Z^*_i is non-decreasing, this in turn implies that $r^*_{0,i}$ is non-decreasing. When combined with Lemmas 5, 6 and 11, this establishes the first statement of the lemma.

Combining the relation for $\bar{a}$ in (33) with the definition (23) of $r_{0,i}$, the definition of z^* and $r^*_{j,i} \circ Z^*_i$ given in (20) and the time-reversed version of (12), we see that for a.e. $t \geq 0$ and $i \in \mathcal{J}$,

$$\bar{a}^*_i(t) \;=\; z^*_i(-t) - \sum_{j\in\mathcal{J}} r^*_{j,i}(Z^*_i(-t)) z^*_i(-t) = \bar{\xi}_i(N^*_{-t}) - \sum_{j\in\mathcal{J}} \bar{\gamma}_{i,j}\zeta_j(N^*_{-t}) = \bar{\alpha}_i.$$

Next, for $i \in N^*_{-t}$, we have $b_i(t) = 1$ and hence the second equation in (33), along with the definition of x^* in (20) and the time-reversed version of (13) shows that

$$\bar{s}^*_i(\bar{B}^*_i(t)) \;=\; x^*_i(-t) = \bar{\zeta}_i(N^*_{-t}) = \mu_i.$$

For $i \notin N^*_{-t}$, we know that for a.e. t, $Q^*_i(-t) = 0$ and hence $q^*_i(-t) = 0$ (since Q^* is absolutely continuous). In turn, due to the expressions for q^* and s^* in (20), this implies that $x^*_i(t) = z^*_i(t)$ and $\bar{\xi}_i(N^*_{-t}) = \bar{\zeta}_i(N^*_{-t})$. Consequently, using the second and sixth equations in (33), we obtain

$$\bar{s}^*_i(\bar{B}^*_i(t)) = \frac{x^*_i(-t)}{b^*_i(-t)} \;=\; \frac{z^*_i(-t)}{b^*_i(-t)} = s^*_i(B^*_i(-t)) = \mu_i \bar{\xi}_i(N^*_{-t})/\bar{\zeta}_i(N^*_{-t}) = \mu_i.$$

Thus the second equation of the lemma holds for all $i \in \mathcal{J}$.

Finally, for a.e. $t \geq 0$ and $i, j \leq n$, using the third and sixth relations in (33), and the relations for x^*, z^* and r^* in (20), we conclude that

$$\bar{r}^*_{i,j}(\bar{Z}^*_j(t)) \quad = \quad r^*_{j,i}(Z_i(-t))z^*_i(-t)/x^*_j(-t) = \bar{\gamma}_{i,j},$$

thus completing the proof. ■

5 Large deviations

In this section we take advantage of the well known sample path large deviation principle for the sequence $(\Gamma_k(\tilde{S}, \tilde{A}, \tilde{R}))_{k \in \mathbb{N}}$ and the partial functional continuity of the set-valued mapping $\mathcal{Q}$ in order to derive large deviation bounds for sequence $(\Gamma_k(\tilde{S}, \tilde{A}, \tilde{R}, \tilde{X}, \tilde{B}, \tilde{Y}, \tilde{Z}, \tilde{Q}))_{k \in \mathbb{N}}$.

We first recall some basic elements of large deviations theory [9]. An $\mathbb{R}_+ \cup \{\infty\}$-valued and lower semicontinuous functions is called a *rate function.* A rate function is *good* if it has compact level sets. We will say that a sequence of random paths $(\tilde{D}^k)_{k \in \mathbb{N}}$ in $\mathcal{D}_c$ satisfies a *sample path large deviation principle* with rate function $J \colon \mathcal{D}_c \to \mathbb{R}_+ \cup \{\infty\}$ if for every measurable set $A \subset \mathcal{D}_c$

$$\limsup_{k \to \infty} \frac{1}{k} \log P(\tilde{D}^k \in A) \quad \leq \quad - \inf_{D \in A^c} J(D)$$

and

$$\liminf_{k \to \infty} \frac{1}{k} \log P(\tilde{D}^k \in A) \quad \geq \quad - \inf_{D \in A^o} J(D),$$

where A^c is the closure and A^o the interior of A with respect to the topology induced by the norm (1). Sample path large deviation principles on product path spaces are defined analogously.

In Section 5.1 we state the large deviation principle satisfied by the sequence of scaled Jackson network primitives and also introduce a corresponding "pseudo local rate function". In Section 5.3 we introduce a time-reversed version of this local rate function. In Section 5.2, we use a type of contraction principle established in [27] to provide a variational characterization for asymptotic probabilities related to the sequence of scaled Jackson network processes $(\tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B})$. In Section 5.4 we use time-reversal arguments to explicitly find the solution to this variational problem for the particular event of a large queue build up at time 0. Finally, these elements are combined to produce the proof of the main theorem in Section 5.5.

5.1 LDP for Jackson Network Primitives

We first introduce some basic notation that is needed in order to state the LDP for the sequence of scaled Jackson network primitives. For $\delta \geq 0$, we define the function $\Phi_\delta\colon \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ by

$$\Phi_\delta(\delta') \quad = \quad \begin{cases} \delta - \delta' & \text{if } \delta' = 0, \\ \infty & \text{if } \delta = 0 \text{ and } \delta' > 0, \\ \delta - \delta' + \delta' \log \frac{\delta'}{\delta} & \text{otherwise.} \end{cases}$$

For a vector π in $\mathbb{R}^n_+$ satisfying

$$\pi_0 \quad := \quad 1 - \sum_{i \in \mathcal{J}} \pi_i \geq 0,$$

we define the function $\Psi_\pi\colon \mathbb{R}^n_+ \to \mathbb{R}_+ \cup \{\infty\}$ by

$$\begin{aligned} \Psi_\pi(\pi') \quad &:= \quad \begin{cases} \sum_{i=0}^n \pi'_i \log \frac{\pi'_i}{\pi_i} & \text{if } \pi'_0 := 1 - \sum_{i \in \mathcal{J}} \pi'_i \geq 0, \\ \infty & \text{otherwise} \end{cases} \\ &= \quad \Phi_{1 - \sum_{i \in \mathcal{J}} \pi_i}\left(1 - \sum_{i \in \mathcal{J}} \pi'_i\right) + \sum_{i \in \mathcal{J}} \Phi_{\pi_i}(\pi'_i). \end{aligned}$$

The sequence of scaled Jackson network primitives $(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}))_{k \in \mathbb{N}}$ satisfies a sample path large deviation principle [26] with good rate function $I\colon \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma \to \mathbb{R}_+ \cup \{\infty\}$ given by $I(A, S, R) = \infty$ if $S(0) \neq 0$, $A(0) \neq 0$, $R(0) \neq 0$ or one component of these network primitives is not absolutely continuous, and

$$I(A, S, R) \quad := \quad \sum_{i \in \mathcal{J}} \int_{\mathbb{R}} \left(\Phi_{\alpha_i}(a_i(t)) + \Phi_{\mu_i}(s_i(t)) + \Psi_{\gamma_i}(r_i(t))\right) dt,$$

otherwise. Here, γ_i (resp. r_i) denotes the ith column vector of the matrix γ (resp. r).

The rate function admits a more convenient representation, involving the pseudo-local rate function $L_{\alpha,\mu,\gamma}$, which we define for $\mu', \alpha' \in \mathbb{R}^n_+$, $\gamma' \in \mathbb{R}^n_+$ and $\beta' \in [0,1]^n$, by

$$L_{\alpha,\mu,\gamma}(\alpha', \mu', \gamma', \beta') \quad := \quad \sum_{i \in \mathcal{J}} \left(\Phi_{\alpha_i}(\alpha'_i) + \beta'_i \Phi_{\mu_i}(\mu'_i) + \beta'_i \mu'_i \Psi_{\gamma_i}(\gamma'_i)\right).$$

In addition, for absolutely continuous primitive paths (A, S, R), with $S(0) = 0$, $A(0) = 0$ and $R(0) = 0$ and a network path $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ we can use a simple change of variables to rewrite the rate function $I(A, S, R)$ in the form

$$I(A, S, R) \;=\; \int_{\mathbb{R}} L_{\alpha,\mu,\gamma}\left(a(t), s(B(t)), r(Z(t)), b(t)\right) dt$$

where for a matrix of functions $M = (M_{i,j})_{i,j\in\mathcal{J}}$ and a vector $v = (v_j)_{j\in\mathcal{J}}$ we define the matrix $M(v)$ by $(M(v))_{i,j} = M_{i,j}(v_j)$.

For later use, we prove an elementary lemma about the zeros of the "pseudo" local rate function $L_{\alpha,\mu,\gamma}$.

Lemma 13 $L_{\alpha,\mu,\gamma} \geq 0$. *Moreover, if (A, S, R) are absolutely continuous network primitive paths and $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ then for a.e. $t \in \mathbb{R}$ we have*

$$L_{\alpha,\mu,\gamma}(a(t), s(B(t)), r(Z(t)), b(t)) \;=\; 0,$$

if and only if $a(t) = \alpha$, $s(B(t)) = \mu$ and $r(Z(t)) = \gamma$.

Proof. Since (A, S, R) are absolutely continuous, by Remark 3(b), $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ are also absolutely continuous and their derivatives satisfy the set of equations (10) for a.e. t. Fix a $t \in \mathbb{R}$ such that the equations (10) are satisfied and define $\alpha' := a(t)$, $\mu' := s(B(t))$, $\gamma' = r(Z(t))$, $\beta' = b(t)$ and $N := \{i \in \mathcal{J} \colon Q_i(t) > 0\}$. Since α_i is the only zero point of Φ_{α_i} it immediately follows that $\alpha'_i = \alpha_i$ for every $i \in \mathcal{J}$. Now, define $M := \{i \in \mathcal{J} : \mu'_i\beta'_i = 0\}$. Then a similar argument (and the fact that clearly $\beta'_i \neq 0$ for $i \notin M$) shows that $\mu'_i = \mu_i$ and $\gamma_i = \gamma'_i$ for $i \notin M$. Thus to prove the lemma, it suffices to show that $M = \emptyset$.

From the fourth and fifth relations in (10), first notice that $\beta'_i = 1$ for $i \in N$. By the same argument just invoked above, this implies that $\mu'_i = \mu_i$ for $i \in N$. In turn, since $\mu > 0$, this implies $M \cap N = \emptyset$. Next, observe that the first, second, third and sixth relations in (10), when combined with the fact that $\alpha = \alpha'$ and the definition of M, show that for $i \notin N$,

$$\mu'_i\beta'_i \;=\; \alpha'_i + \sum_{j\in\mathcal{J}} \gamma'_{i,j}\mu'_j\beta'_j = \alpha_i + \sum_{j\in\mathcal{J}\setminus M} \gamma'_{i,j}\mu'_j\beta'_j.$$

Now suppose M is non-empty. Then applying the last equation for $i \in M \subseteq \mathcal{J} \setminus N$, the left-hand side is zero and so we conclude that $\alpha_i = 0$ and $\gamma'_{i,j} = 0$ for all $i \in M$ and $j \notin M$. However, this contradicts the fact that

$$\sum_{k=0}^{\infty} \gamma^k \alpha = (I - \gamma)^{-1} \alpha = \lambda > 0,$$

and thus proves that $M = \emptyset$, concluding the proof of the lemma. ■

5.2 Variational Representation

If for every primitive path $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$, there existed a unique network path (X, Z, Q, Y, B) in $\mathcal{Q}(A, S, R)$ and the map from (A, S, R) to (X, Z, Q, Y, B) were continuous, then we could use the so-called contraction principle [9, 16] to obtain a large deviation principle for the scaled sequence of network processes with a rate function given explicitly in terms of the rate function I for the primitive processes. Although, we do not have precisely this situation here, we can use the upper-semicontinuity of the map $\mathcal{Q}$ to obtain a contraction-like principle.

Theorem 14 *For every measurable set $M \subset \mathcal{M}$*

$$\begin{aligned} \limsup_{k\to\infty} \frac{1}{k} \log P\left(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \in M\right) \\ \leq -\inf_{(S,A,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma,\ \{(A,S,R)\}\times\mathcal{Q}(A,S,R)\cap M\neq\emptyset} I(A, S, R), \quad (35) \\ \liminf_{k\to\infty} \frac{1}{k} \log P\left(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \in M\right) \\ \geq -\inf_{(S,A,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma,\ \{(A,S,R)\}\times\mathcal{Q}(A,S,R)\subset M^o} I(A, S, R), \quad (36) \end{aligned}$$

Proof. From Theorem 2 in [27], we know that the mapping $\mathcal{I}_\alpha \times (\mathcal{I}_\mu \cap \mathcal{C}_\mu) \times \mathcal{I}_\gamma \ni (A, S, R) \mapsto \{(A, S, R)\} \times \mathcal{Q}(A, S, R)$ is partially upper semicontinuous in the sense that if $(A_k, S_k, R_k)_{k\in\mathbb{N}}$ is a convergent sequence in $\mathcal{I}_\alpha \times (\mathcal{I}_\mu \cap \mathcal{C}_\mu) \times \mathcal{I}_\gamma$ with a continuous limit (A, S, R) and $(X_k, Z_k, Q_k, Y_k, B_k) \in \mathcal{Q}(A_k, S_k, R_k)$ for every $k \in \mathbb{N}$, then there exists a subsequence of $(X_k, Z_k, Q_k, Y_k, B_k)_{k\in\mathbb{N}}$ which converges and each such convergent subsequence converges to an element of $\mathcal{Q}(A, S, R)$.

For $i \in \mathcal{J}$ we define the processes $\hat{S}_i$ and $\hat{Z}_i$ by linearly interpolating the process $\tilde{S}_i$, respectively, $\tilde{Z}_i$ between jumps. Furthermore we define $\hat{Q} := \tilde{X} - \hat{Z}$. Using the properties of Theorem 2 one can verify $(\tilde{X}, \hat{Z}, \hat{Q}, \tilde{Y}, \tilde{B}) \in \mathcal{Q}(\tilde{A}, \hat{S}, \tilde{R})$ with probability 1. Note that the modified process $\hat{S}$ is continuous and therefore the modified network primitives $(\tilde{A}, \hat{S}, \tilde{R})$ fall into the range of the indicated upper semicontinuity.

Using the mentioned upper semicontinuity, one can prove bounds (35) and (36) with $(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B})$ replaced by $(\tilde{A}, \hat{S}, \tilde{R}, \tilde{X}, \hat{Z}, \hat{Q}, \tilde{Y}, \tilde{B})$ for measurable sets $M \subset \mathcal{I}_\alpha \times (\mathcal{I}_\mu \cap \mathcal{C}_\mu) \times \mathcal{I}_\gamma \times \mathcal{I}_\lambda \times \mathcal{I}_\lambda \times \mathcal{D}_0 \times \mathcal{I}_{e-\varrho} \times \mathcal{I}_\varrho$ completely analogous as in the proof of Theorem 3 of [26].

By mimicking standard results of large deviations theory (e.g. the proof of statement (a) in Lemma 4.1.5 of [9]), one can extend these bounds to be valid for all measurable sets M of the larger space $\mathcal{M}$.

The statement of the theorem now follows from the fact that sequences of processes $(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}))_{k\in\mathbb{N}}$ and $(\Gamma_k(\tilde{A}, \hat{S}, \tilde{R}, \tilde{X}, \hat{Z}, \hat{Q}, \tilde{Y}, \tilde{B}))_{k\in\mathbb{N}}$ are exponentially equivalent [9], since the distance of the kth elements of these sequences is at most $1/k$. ■

We will be interested in the particular case when M is the event that the queue length at 0 is greater than a prescribed level Q^0. We will solve the above variational problem for this particular choice of M in Section 5.4. The solution of this variational problem will be facilitated by a time-reversal argument, which will make use of a time-reversed version of the pseudo local rate function, which we introduce in the next section.

5.3 Time-reversal of the pseudo local rate function

Our objective in this section is to identify a time-reversed cost function $\bar{I}$ on paths in the time-reversed network such that the minimum I-cost of inputs (A, S, R) for which there exists a corresponding Q that reaches Q^0 at time 0 in the forward network is equal to the minimum $\bar{I}$-cost over all inputs $(\bar{A}, \bar{S}, \bar{R})$ for which there exists a corresponding $\bar{Q}$ that drains from Q^0 to 0 in the time-reversed network. This transforms the large deviations optimization problem of finding the I-optimal trajectory that goes from 0 to Q^0 (in the forward network) into a fluid optimization problem of finding the $\bar{I}$-optimal trajectory that goes from Q^0 to 0 (in the time-reversed network). The fact that the latter class of optimization problems if often easier to solve provides the motivation for trying to establish such an equivalence [16]. Indeed, in the proof of Theorem 15 we establish such an equivalence for the Jackson network, and show that the minimizing trajectories can be identified easily as a consequence. A major ingredient in this reformulation is the proof of the local equivalence between forward and backward trajectories proved in Theorem 15.

Theorem 15 *Given any absolutely continuous network primitive paths* (A, S, R) *and network paths* $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ *that are reversible, and corresponding time reversed quantities* $(\bar{A}, \bar{S}, \bar{R})$ *and* $(\bar{X}, \bar{Z}, \bar{Q}, \bar{Y}, \bar{B})$*, we have for a.e.* $t \in \mathbb{R}$*,*

$$\begin{aligned} & L_{\bar{\alpha},\mu,\bar{\gamma}}(\bar{a}(-t), \bar{s}(\bar{B}(-t)), \bar{r}(\bar{Z}(-t)), \bar{b}(-t)) \\ &\quad = L_{\alpha,\mu,\gamma}(a(t), s(B(t)), r(Z(t)), b(t)) - \sum_{i\in\mathcal{J}} q_i(t) \log \frac{\mu_i}{\lambda_i}, \end{aligned}$$

where $L_{\bar{\alpha},\mu,\bar{\gamma}}$ *is defined as* $L_{\alpha,\mu,\gamma}$ *but with the time-reversed mean rates.*

Proof. We first derive a few identities that are required for the proof. The definition (14) of $\bar{\alpha}$ along with the traffic equation (3) yields the first identity

$$\sum_{i\in\mathcal{J}} (\bar{\alpha}_i - \alpha_i) \;=\; \sum_{i\in\mathcal{J}} (\gamma_{0,i}\lambda_i - \alpha_i) = 0. \tag{37}$$

Now fix $t \in \mathbb{R}$ for which the network equations in (10) and (33) hold. In what follows, for $i, j \in \{1, \dots, n\}$ the functions s_i and $r_{i,j}$ are evaluated at $B_i(t)$ and $Z_j(t)$ respectively, while the functions a, x, y, b, z and q are evaluated at t. Analogously, the functions $\bar{s}_i$ and $\bar{r}_{i,j}$ are evaluated at $\bar{B}_i(-t)$ and $\bar{Z}_j(-t)$ respectively, while $\bar{a}$, $\bar{x}$, $\bar{y}$, $\bar{b}$, $\bar{z}$ and $\bar{q}$ are evaluated at $-t$. For notational conciseness we will omit the arguments of the functions.

Define $N := \{i \in \mathcal{J} \colon Q_i > 0\}$ and $N_c := \mathcal{J} \setminus N$. Then $i \in N$ implies $b_i = 1$ and $i \in N_c$ implies $q_i = 0$. This leads to the second identity,

$$\sum_{i\in\mathcal{J}} a_i - \sum_{i\in\mathcal{J}} r_{0,i} z_i \;=\; \sum_{i\in N} (x_i - s_i) = \sum_{i\in N} (x_i - s_i b_i)\,, \tag{38}$$

which simply states that the net external arrival rate into the network minus the net departure rate from the network should equal the net rate of increase of the sum of all queues in the network.

Next, we observe that $L_{\bar{\alpha},\mu,\bar{\gamma}}(\bar{a},\bar{s},\bar{r},\bar{b})$ is equal to

$$\sum_{i\in\mathcal{J}}\Bigg(\bar{\alpha}_i-\bar{a}_i+\bar{a}_i\log\frac{\bar{a}_i}{\bar{\alpha}_i}+\bar{b}_i\mu_i-\bar{b}_i\bar{s}_i+\bar{b}_i\bar{s}_i\log\frac{\bar{s}_i}{\mu_i}$$
$$+\sum_{j\in\mathcal{J}}\bar{z}_i\bar{r}_{j,i}\log\frac{\bar{r}_{j,i}}{\bar{\gamma}_{j,i}}+\bar{z}_i\bar{r}_{0,i}\log\frac{\bar{r}_{0,i}}{\bar{\gamma}_{0,i}}\Bigg).$$

Adding and subtracting $\sum_{i\in\mathcal{J}}\alpha_i$, using (37) and rewriting $\mu,\bar{\gamma},\bar{a},\bar{s},\bar{r},\bar{b}$ in terms of the corresponding forward network quantities as specified in Remark 10, the above display can be shown to be equal to

$$\sum_{i\in\mathcal{J}}\left(\alpha_i-r_{0,i}z_i+r_{0,i}z_i\log\left(\frac{r_{0,i}z_i}{\gamma_{0,i}\lambda_i}\right)+b_i\mu_i\right)-\sum_{i\in N_c}b_is_i-\sum_{i\in N}x_i$$
$$+\sum_{i\in N_c}b_is_i\log\frac{s_i}{\mu_i}+\sum_{i\in N}x_i\log\frac{x_i}{\mu_i}+\sum_{i,j\in\mathcal{J}}r_{i,j}z_j\log\left(\frac{r_{i,j}z_j\lambda_i}{\bar{z}_i\gamma_{i,j}\lambda_j}\right)+\sum_{i\in\mathcal{J}}a_i\log\left(\frac{a_i\lambda_i}{x_i\alpha_i}\right).$$

Adding and subtracting $\sum_{i=1}^n a_i$ and rearranging terms, the above display can be rewritten as

$$\sum_{i\in\mathcal{J}}(\alpha_i-a_i+b_i\mu_i-b_is_i)+\sum_{i,j\in\mathcal{J}}r_{i,j}z_j\log\frac{r_{i,j}}{\gamma_{i,j}}+\sum_{i\in\mathcal{J}}r_{0,i}z_i\log\frac{r_{0,i}}{\gamma_{0,i}}+\sum_{i\in\mathcal{J}}a_i\log\frac{a_i}{\alpha_i}$$
$$+\left(\sum_{i\in\mathcal{J}}a_i+\sum_{i\in N}(b_is_i-x_i)-\sum_{i\in\mathcal{J}}r_{0i}z_i\right)+\sum_{i\in N_c}b_is_i\log\frac{s_i}{\mu_i}+\sum_{i\in N}x_i\log\frac{x_i}{\mu_i}$$
$$+\sum_{i,j\in\mathcal{J}}r_{i,j}z_j\log\left(\frac{z_j\lambda_i}{x_i\lambda_j}\right)+\sum_{i\in\mathcal{J}}r_{0,i}z_i\log\frac{z_i}{\lambda_i}+\sum_{i\in\mathcal{J}}a_i\log\frac{\lambda_i}{x_i}.$$

Using (38), the definition of $L_{\alpha,\mu,\gamma}$ and the fact that $z_j=x_j$ for $j\in N_c$, $z_j=s_j$ for $j\in N$, and $b_i=1$ for $i\in N$ this reduces to

$$L_{\alpha,\mu,\gamma}(a,s,r,b)-\sum_{i\in N}s_i\log\frac{s_i}{\mu_i}+\sum_{i\in N}x_i\log\frac{x_i}{\mu_i}+\sum_{j\in N_c}\sum_{i\in\mathcal{J}}r_{i,j}x_j\log\frac{x_j}{\lambda_j}$$
$$+\sum_{j\in N}\sum_{i\in\mathcal{J}}r_{i,j}s_j\log\frac{s_j}{\lambda_j}+\sum_{i\in N_c}r_{0,i}x_i\log\frac{x_i}{\lambda_i}+\sum_{i\in N}r_{0,i}s_i\log\frac{s_i}{\lambda_i}$$
$$+\sum_{i,j\in\mathcal{J}}r_{i,j}z_j\log\frac{\lambda_i}{x_i}+\sum_{i\in\mathcal{J}}a_i\log\frac{\lambda_i}{x_i}.$$

By Remark 3(b), $q_i = x_i - s_i$ for $i \in N$, and so a rearrangement of terms in the last display yields

$$L_{\alpha,\mu,\gamma}(a,s,r,b) + \sum_{i\in N} q_i \log \frac{\lambda_i}{\mu_i} + \sum_{i\in N} \left(r_{0,i}s_i - s_i + \sum_{j\in\mathcal{J}} r_{j,i}s_i \right) \log \frac{s_i}{\lambda_i}$$
$$+ \sum_{i\in\mathcal{J}} \left(a_i + \sum_{j\in\mathcal{J}} r_{i,j}z_j - \sum_{i\in N_c} r_{0,i}x_i - \sum_{i\in N} x_i - \sum_{i\in N_c}\sum_{j\in\mathcal{J}} r_{j,i}x_i \right) \log \frac{\lambda_i}{x_i}.$$

Observing that the last two terms are equal to zero due to the definition (23) of $r_{0,i}$ and the first equation in (10), the last five displays can be combined to obtain

$$\begin{aligned} L_{\bar{\alpha},\mu,\bar{\gamma}}(\bar{a},\bar{s},\bar{r},\bar{b}) &= L_{\alpha,\mu,\gamma}(a,s,r,b) + \sum_{i\in N} q_i \log \frac{\lambda_i}{\mu_i} \\ &= L_{\alpha,\mu,\gamma}(a,s,r,b) - \sum_{i\in\mathcal{J}} q_i \log \frac{\mu_i}{\lambda_i}, \end{aligned}$$

where the last equation follows because $q_i = 0$ for $i \in N_c$. This proves the theorem. ∎

5.4 Solution of a Variational Problem

We now solve the variational problem introduced in Section 5.2 for the particular event of the queue reaching a level Q^0 at time 0. Indeed, the following is the main result of this section.

Proposition 16 *Only the network primitives* (A^*, S^*, R^*) *and network behavior* $(X^*, Z^*, Q^*, Y^*, B^*)$ *satisfy* $Q^*(0) \geq Q^0$ *and*

$$I(A^*,S^*,R^*) = \inf_{\substack{(A,S,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma \text{ such that} \\ (X,Z,Q,B,Y)\in\mathcal{Q}(A,S,R) \\ \text{exists with } Q(0)\geq Q^0}} I(A,S,R). \tag{39}$$

The remainder of this section is devoted to the proof of this proposition. We first establish two preliminary lemmas.

Lemma 17 *Suppose an absolutely continuous $R \in \mathcal{I}_\gamma$ satisfies $R(0) = 0$ and for a.e. $c \in \mathbb{R}_+$ and $j \in \mathcal{J}$ there exists $N_c \subseteq \mathcal{J}$ such that*

$$r_{i,j}(c) \;=\; \bar{\gamma}_{j,i}\bar{\zeta}_i(N_c)/\bar{\xi}_j(N_c) \quad \textit{for } i \in \mathcal{J}. \tag{40}$$

Then $v_i \le \sum_{j\in\mathcal{J}} R_{i,j}(v_j)$ implies $v \le 0$ for every $v \in \mathbb{R}^n$.

Proof. Suppose there exists $v \in \mathbb{R}^n$ such that $v_i \le \sum_{j\in\mathcal{J}} R_{i,j}(v_j)$ for $i \in \mathcal{J}$ (otherwise the lemma is vacuously true) and let $M := \{i \in \mathcal{J} \colon v_i > 0\}$. We will argue by contradiction to show that $M = \emptyset$. Indeed, suppose $M \neq \emptyset$. Then condition (40), along with the time-reversed version of (12), implies

$$\sum_{i\in\mathcal{J}} r_{i,j}(c) = \sum_{i\in\mathcal{J}} \frac{\bar{\gamma}_{j,i}\bar{\zeta}_i(N_c)}{\bar{\xi}_j(N_c)} = 1 - \frac{\bar{\xi}_j(N_c) - \sum_{i\in\mathcal{J}} \bar{\gamma}_{j,i}\bar{\zeta}_i(N_c)}{\bar{\xi}_j(N_c)} = 1 - \frac{\bar{\alpha}_j}{\bar{\xi}_j(N_c)}.$$

for every $j \in \mathcal{J}$ and a.e. $c \ge 0$. Hence, there exists $\varepsilon > 0$ such that

$$\sum_{i\in\mathcal{J}} R_{i,j}(c) \;\le\; c(1 - \varepsilon\bar{\alpha}_j)$$

for every $c \in \mathbb{R}_+$ and $j \in \mathcal{J}$. As a result, we have

$$\sum_{i\in M} v_i \le \sum_{i\in M}\sum_{j\in\mathcal{J}} R_{i,j}(v_j) = \sum_{j\in\mathcal{J}}\sum_{i\in M} R_{i,j}(v_j) \le \sum_{j\in M}\sum_{i\in\mathcal{J}} R_{i,j}(v_j) \le \sum_{j\in M} v_j(1-\varepsilon\bar{\alpha}_j),$$

where the second inequality uses the fact that $R_{i,j}(v_j) \ge 0$ for all $j \in M$ and $R_{i,j}(v_j) \le 0$ for $j \notin M$. Since $\varepsilon > 0$ and $\bar{\alpha} \ge 0$, this holds only if $\bar{\alpha}_j = 0$ for every $j \in M$. However, because $\sum_{k=0}^{\infty} \bar{\gamma}^k\bar{\alpha} = (I - \bar{\gamma})^{-1}\bar{\alpha} = \lambda > 0$ there is a $j' \in M$ and $i' \in \mathcal{J} \setminus M$ such that $\bar{\gamma}_{j',i'} > 0$. Then condition (40) implies $R_{i',j'}(v_{j'}) > 0$, which leads to the contradiction (compare previous display)

$$\sum_{i\in M} v_i \;\le\; \sum_{j\in\mathcal{J}}\sum_{i\in M} R_{i,j}(v_j) \le \sum_{j\in M}\sum_{i\in\mathcal{J}} R_{i,j}(v_j) - R_{i',j'}(v_{j'}) \le \sum_{j\in M} v_j - R_{i',j'}(v_{j'}).$$

This shows that $M = \emptyset$ and concludes the proof. ∎

Lemma 18 *If* $(A, S, R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$, $(X, Z, Q, Y, B) \in \mathcal{Q}(A, S, R)$ *and* $I(A, S, R) < \infty$, *then* (A, S, R)*and* (X, Z, Q, Y, B) *are reversible. If, in addition,* $Q(0) = Q^0$, *time-reversed quantities are marked with a bar,*

$$\begin{aligned} 0 &= \int_0^\infty L_{\alpha,\mu,\gamma}(a(t), s(B(t)), r(Z(t)), b(t))\, dt \\ &= \int_{-\infty}^0 L_{\bar{\alpha},\mu,\bar{\gamma}}(\bar{a}(t), \bar{s}(\bar{B}(t)), \bar{r}(\bar{Z}(t)), \bar{b}(t))\, dt, \end{aligned}$$

then $(A, S, R, X, Z, Q, Y, B) = (A^*, S^*, R^*, X^*, Z^*, Q^*, Y^*, B^*)$.

Proof. The first statement follows from statement 5 in Lemma 17 of [27] and Lemma 11. For the second statement, first note that by Lemma 13, we must have, for $i \in \mathcal{J}$ and a.e. $t \geq 0$: $s(B(t)) = \mu$, $a(t) = \alpha$ and $r_i(Z_i(t)) = \gamma_i$ and $\bar{s}(\bar{B}(-t)) = \mu$, $\bar{a}(-t) = \bar{\alpha}$ and $\bar{r}_i(\bar{Z}_i(-t)) = \bar{\gamma}_i$. As usual, we define $N_t := \{i \in \mathcal{J}\colon Q_i(t) > 0\}$ for $t \in \mathbb{R}$. Now, the last four equations in (10) along with the definitions (12) and (13) show that $b_i(t) = 1$ if $i \notin N_t$ and $x_i(t) = z_i(t)$ if $i \notin N_t$. Combining this with all the equations in (10), it is straightforward to deduce that for $i \in \mathcal{J}$ and a.e. $t \geq 0$,

$$\begin{aligned} x_i(t) &= \xi_i(N_t), & z_i(t) &= \zeta_i(N_t), \\ q_i(t) &= \xi_i(N_t) - \zeta_i(N_t), & y_i(t) &= 1 - \zeta_i(N_t)/\mu_i, \\ b_i(t) &= \zeta_i(N_t)/\mu_i. \end{aligned}$$

Analogously, one can use Remark 10 to deduce that for a.e. $t \leq 0$ and $i, j \in \mathcal{J}$,

$$\begin{aligned} a_i(t) &= \bar{\gamma}_{0,i}\bar{\zeta}_i(N_t), & s_i(B_i(t)) &= \mu_i\bar{\xi}_i(N_t)/\bar{\zeta}_i(N_t), \\ r_{i,j}(Z_j(t)) &= \bar{\gamma}_{j,i}\bar{\zeta}_i(N_t), & x_i(t) &= \bar{\zeta}_i(N_t) - \bar{\xi}_i(N_t), \\ z_i(t) &= \bar{\xi}_i(N_t), & q_i(t) &= \bar{\zeta}_i(N_t) - \bar{\xi}_i(N_t), \\ y_i(t) &= 1 - \bar{\zeta}_i(N_t)/\mu_i, & b_i(t) &= \bar{\zeta}_i(N_t)/\mu_i. \end{aligned}$$

Now, since $Q = X - Z \geq 0$ and $A(0) = 0$, we have for $i \in \mathcal{J}$

$$Z_i(0) \;\leq\; X_i(0) = \sum_{j \in \mathcal{J}} R_{i,j}(Z_j(0)).$$

Therefore $Z(0) \leq 0$ by Lemma 17. Since $R(0) = 0$, we conclude that $R_{i,j}(c) = \gamma_{i,j}c$ for $c \geq Z_j(0)$ and $i, j \in \mathcal{J}$. This implies for $i \in \mathcal{J}$

$$Z_i(0) + Q_i^0 = X_i(0) = \sum_{j \in \mathcal{J}} \gamma_{i,j} Z_j(0)$$

and hence

$$\begin{array}{rclcrcl} Z(0) & = & -(I-\gamma)^{-1}Q^0, & \quad & B(0) & = & \mathrm{diag}(\mu)^{-1}Z(0), \\ Y(0) & = & -B(0), & & X(0) & = & -\gamma(I-\gamma)^{-1}Q^0. \end{array}$$

The above assertions, when combined, imply the statement of the lemma. ■

We are now in a position to complete the proof of Proposition 16.

Proof of Proposition 16. If $(A,S,R) \in \mathcal{I}_\alpha \times \mathcal{I}_\mu \times \mathcal{I}_\gamma$ are network primitives with $I(A,S,R) < \infty$ and $(X,Z,Q,Y,B) \in \mathcal{Q}(A,S,R)$ satisfies $Q(0) \geq Q^0$ then this pair is reversible by Lemma 18. We let $(\bar{A},\bar{S},\bar{R})$ and $(\bar{X},\bar{Z},\bar{Q},\bar{Y},\bar{B})$ be the corresponding time reversed paths. From statement 6 of Lemma 17 in [27], we know $\lim_{|t|\to\infty} Q(t) = 0$. Hence we can deduce from Theorem 15 that

$$\begin{aligned} I(A,S,R) &= \int_{-\infty}^{\infty} L_{\alpha,\mu,\gamma}(a(t), s(B(t), r(Z(t)), b(t))\, dt \\ &= \int_{-\infty}^{0} L_{\bar{\alpha},\mu,\bar{\gamma}}(\bar{a}(-t), \bar{s}(\bar{B}(-t)), \bar{r}(\bar{Z}(-t)), \bar{b}(-t))\, dt \\ &\quad + \int_{-\infty}^{0} \sum_{i\in\mathcal{J}} q_i(t) \log \frac{\mu_i}{\lambda_i} \\ &\quad + \int_{0}^{\infty} L_{\alpha,\mu,\gamma}(a(t), s(B(t)), r(Z(t)), b(t))\, dt \geq \sum_{i\in\mathcal{J}} Q_i^0 \log \frac{\mu_i}{\lambda_i}. \end{aligned}$$

By Proposition 18, equality in the last display is equivalent to $(A,S,R,X,Z,Q,Y,B) = (A^*,S^*,R^*,X^*,Z^*,Q^*,Y^*,B^*)$. ■

5.5 Proof of the Main Result

Proof of Theorem 7. By Lemma 6, Proposition 16, Proposition 2(1) and a simple scaling argument, it follows that for every $c > 1$, the network primitives $(\Gamma_{1/c}A^*, \Gamma_{1/c}S^*, \Gamma_{1/c}R^*)$ give rise to the unique quintuple of network paths $(\Gamma_{1/c}X^*, \Gamma_{1/c}Z^*, \Gamma_{1/c}Q^*, \Gamma_{1/c}Y^*, \Gamma_{1/c}B^*) \in \Lambda_{\alpha,\mu,\gamma}$ and satisfy $I(\Gamma_{1/c}A^*, \Gamma_{1/c}S^*, \Gamma_{1/c}R^*) = cI(A^*,S^*,R^*)$ and $\Gamma_{1/c}Q(0) > Q^0$. Hence the lower bound (36) of Theorem 14 yields, for the open set

$$M := \{(A,S,R,X,Z,Q,Y,B) \in \mathcal{M}\colon Q(0) > Q^0\},$$

the estimate

$$\begin{aligned}
&\liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}\left(\tilde{Q}(0) \geq kQ^0\right) \\
&\qquad \geq \liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}\left(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \in M\right) \\
&\qquad \geq -\inf_{(A,S,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma,\ \{(A,S,R)\}\times\mathcal{Q}(A,S,R)\subset M} I(A,S,R) \\
&\qquad \geq -\lim_{c\downarrow 1} I(\Gamma_{1/c}A^*, \Gamma_{1/c}S^*, \Gamma_{1/c}R^*) = -I(A^*, S^*, R^*). \qquad (41)
\end{aligned}$$

On the other hand, for any fixed $\varepsilon > 0$, since the network primitives (A^*, S^*, R^*) are the unique solution to the minimization problem (39) and the rate function I has compact level sets we get from the large deviation upper bound (35) of Theorem 14 applied to the closed set

$$M' := \left\{ \begin{array}{c} (A,S,R,X,Z,Q,Y,B) \in \mathcal{M}\colon Q(0) \geq Q^0, \\ \|(A,S,R,X,Z,Q,Y,B) - (A^*,S^*,R^*,X^*,Z^*,Q^*,Y^*,B^*)\| \geq \varepsilon \end{array} \right\},$$

the estimate

$$\begin{aligned}
&\limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}\Big(\|\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \\
&\qquad\qquad -(A^*,S^*,R^*,X^*,Z^*,Q^*,Y^*,B^*)\| \geq \varepsilon, \tilde{Q}(0) \geq kQ^0\Big) \\
&\quad = \limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}\left(\Gamma_k(\tilde{A}, \tilde{S}, \tilde{R}, \tilde{X}, \tilde{Z}, \tilde{Q}, \tilde{Y}, \tilde{B}) \in M'\right) \\
&\quad \leq -\inf_{(A,S,R)\in\mathcal{I}_\alpha\times\mathcal{I}_\mu\times\mathcal{I}_\gamma,\ \{(A,S,R)\}\times\mathcal{Q}(A,S,R)\cap M'\neq\emptyset} I(A,S,R) \\
&\quad < -I(A^*, S^*, R^*). \qquad (42)
\end{aligned}$$

Bounds (41) and (42) together imply estimate (22), which in turn implies the weak convergence stated in Theorem 7 thus completing its proof. ■

References

[1] V. Anantharam. How large delays build up in a GI/G/1 queue. *Queueing Systems* **5**, 345-368.

[2] V. Anantharam, P. Heidelberger and P. Tsoucas. *Analysis of rare events in continuous time Markov chains via time reversal and fluid approximation.* IBM Research report, #RC 16280, November, 1990.

[3] Rami Atar and Paul Dupuis Large Deviations and Queueing Networks: Methods for Rate Function Identification. *Stoc. Proc. Appl.* **84**, 255-296, 1998.

[4] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis.* Birkhäuser, 1990.

[5] J. Bucklew. *Large Deviations Techniques in Decision, Simulation and Estimation.* Wiley Interscience, New York, 1990.

[6] H. Chen and A. Mandelbaum. Discrete flow networks: bottleneck analysis and fluid approximations. *Math. Oper. Res.* **16**, 2:408-446, 1991.

[7] H. Chen and A. Mandelbaum. Stochastic discrete flow networks: diffusion approximations and bottlenecks. *Ann. Probab.* **19**, 4:1463-1519, 1991.

[8] H. Chen and D.D. Yao. *Fundamentals of Queueing Networks.* Springer-Verlag, Berlin, 2001.

[9] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett, London, 1993.

[10] T. Dieker and M. Mandjes. On asymptotically efficient simulation of large deviation probabilities. *Adv. Appl. Probab.* **37**, 2:539-552, 2005.

[11] P. Dupuis and R. Ellis. Large deviations for Markov processes with discontinuous statistics, II: random walks. *Probab. Theor. Rel. Fields* **91**, 2:153-194, 1992.

[12] P. Dupuis and R. Ellis. *A Weak Convergence Approach to Large Deviations.* John Wiley & Sons, New York, 1997.

[13] P. Dupuis, R. Ellis and A. Weiss. Large deviations for Markov processes with discontinuous statistics I: general upper bounds. *Ann. Probab.* **19**, 3:1280-1297, 1991.

[14] P. Dupuis, H. Ishii and M. Soner. A viscosity solution approach to the asymptotic analysis of queueing systems. *Ann. Probab.* **18**, 1:226-255, 1990.

[15] P. Dupuis and K. Ramanan. Convex duality and the Skorokhod Problem – II. *Probab. Theor. Rel. Fields*, 115:197–236, 1999.

[16] P. Dupuis and K. Ramanan. A time-reversed representation for the tail probabilities of stationary reflected Brownian motion, *Stoch. Proc. Appl.* **98**, 2:253-287, 2002.

[17] P. Dupuis and H. Wang. Importance sampling, large deviations and differential games. *Stoch. Stoch. Rep.* **76**, 481–508, 2004.

[18] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems*, Springer-Verlag, Berlin, 1984.

[19] A. Ganesh and N. O'Connell and D. Wischik. *Big Queues.* Lecture Notes in Mathematics, vol. 1838. Springer-Verlag, Berlin. 2004.

[20] P. Glynn and R. Szechtman. Difficult queuing simulation problems: rare-event simulation for infinite server queues. Proc. of the 34th Winter Simulation Conference, San Diego, California, 416-423, ACM, 2002.

[21] M. Harrison and M. Reiman. Reflected Brownian motion on an orthant. *Ann. Probab.* **9**, 302-308, 1981.

[22] I. Ignatiouk-Robert. Large deviations of Jackson networks. *Ann. Appl. Probab.*, **10**, 3:962–1001, 2000.

[23] J. Jackson. Network of waiting lines. *Oper. Res.* **5**, 518-521, 1957.

[24] F. P. Kelly. *Reversibility and Stochastic Networks.* Wiley, New York, 1979.

[25] K. Majewski. Single class queueing networks with discrete and fluid customers on the time interval $I\!R$, *Queueing Systems* **36** 4:405-435, 2000.

[26] K. Majewski. Large deviation bounds for single class queueing networks and their calculation. *Queueing Systems* **48**, 103-134, 2004.

[27] K. Majewski. Functional continuity and large deviations for the behavior of single class queueing networks. Submitted, October, 2007.

[28] A. Puhalskii. The actional functional for the Jackson networks. *Markov Proc. Rel. Fields* **13**, 1:99–136, 2007.

[29] K. Ramanan and P. Dupuis. Large deviation properties of data sources that share a buffer. *Ann. Appl. Probab.* **8**, 1070-1129, 1998.

[30] J.S. Sadowsky. On the optimality and stability of exponential twisting in Monte Carlo estimation. *IEEE Trans. Inf. Theory* **39**, 119-128, 1993.

[31] A. Shwartz and A. Weiss. Induced rare events: analysis via large deviations and time reversal. *Adv. Appl. Probab.* **25**, 667-689, 1993.

[32] A. Stolyar and K. Ramanan. Largest weighted delay first scheduling: large deviations and optimality. *Ann. Appl. Probab.* **34**, 1:1-27, 2001.

[33] P. Tsoucas. Rare events in series of queues. *Jour. Appl. Probab.* **29** 1:162-175, 1992.