Adv. Appl. Prob. **47,** 402–424 (2015)
Printed in Northern Ireland

COMPUTABLE BOUNDS ON THE SPECTRAL GAP FOR UNRELIABLE JACKSON NETWORKS

PAWEŁ LOREK * ** AND
RYSZARD SZEKLI,* *** *University of Wrocław*

Abstract

The goal of this paper is to identify exponential convergence rates and to find computable bounds for them for Markov processes representing unreliable Jackson networks. First, we use the bounds of Lawler and Sokal (1988) in order to show that, for unreliable Jackson networks, the spectral gap is strictly positive if and only if the spectral gaps for the corresponding coordinate birth and death processes are positive. Next, utilizing some results on birth and death processes, we find bounds on the spectral gap for network processes in terms of the hazard and equilibrium functions of the one-dimensional marginal distributions of the stationary distribution of the network. These distributions must be in this case strongly light-tailed, in the sense that their discrete hazard functions have to be separated from 0. We relate these hazard functions with the corresponding networks' service rate functions using the equilibrium rates of the stationary one-dimensional marginal distributions. We compare the obtained bounds on the spectral gap with some other known bounds.

Keywords: Unreliable Jackson network; spectral gap; exponential ergodicity; Cheeger's constant

2010 Mathematics Subject Classification: Primary 60K25
Secondary 60J25

1. Introduction

We start with a description of the general setting used in this paper. Let $\boldsymbol{X} = (X_t,\ t \geq 0)$ be a Markov process on a countable state space $\mathbb{E}$ with a bounded generator $\boldsymbol{Q}$ and the corresponding semi-group of operators $(P_t,\ t > 0)$ on $L^2(\mathbb{E},\ \pi)$. We assume ergodicity of this process and the existence of the invariant probability measure π. The usual scalar product on $L^2 := L^2(\mathbb{E}, \pi)$ and the corresponding L^2 norm we denote by

$$\langle f, g\rangle_\pi = \sum_{\boldsymbol{n}\in\mathbb{E}} f(\boldsymbol{n})g(\boldsymbol{n})\pi(\boldsymbol{n}), \qquad \|f\|^2 = \langle f, f\rangle_\pi,$$

and by $\mathbf{1}$ the constant function equal to 1 on $\mathbb{E}$. We will use the symbol $\pi(f)$ to denote $\langle f, \mathbf{1}\rangle_\pi = E_\pi(f(X_t))$. We denote the L_2 spectral gap corresponding to $\boldsymbol{X}$ by

$$\operatorname{gap}(\boldsymbol{Q}) := \inf\{-\langle f, \boldsymbol{Q}f\rangle_\pi : \|f\| = 1, \pi(f) = 0\}. \tag{1.1}$$

We say that $\boldsymbol{X} = (X_t,\ t \geq 0)$ has an *exponential rate of convergence* if $\operatorname{gap}(\boldsymbol{Q}) > 0$.

Received 17 June 2010; revision received 25 September 2013.
* Postal address: Mathematical Institute, University of Wroclaw, pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.
** Email address: pawel.lorek@math.uni.wroc.pl
*** Email address: ryszard.szekli@math.uni.wroc.pl
The work of both authors was supported by NCN Research Grant DEC-2011/01/B/ST1/01305.

Then, for reversible processes, the following conditions are equivalent (see, e.g. [10, Theorem 1.9])

(i) for all $f \in L^2(\mathbb{E}, \pi)$,

$$\|P_t f - \pi(f)\| \leq \mathrm{e}^{-\mathrm{gap}(\boldsymbol{Q})t}\|f - \pi(f)\|, \qquad t > 0,$$

(ii) for each $\boldsymbol{e} \in \mathbb{E}$ there exists $C(\boldsymbol{e}) > 0$ such that

$$\|\delta_{\boldsymbol{e}} P_t - \pi\|_{\mathrm{TV}} \leq C(\boldsymbol{e})\mathrm{e}^{-\alpha t}, \qquad t > 0,$$

for some $\alpha > 0$, where $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm, $\delta_{\boldsymbol{e}}$ denotes the measure with single atom at $\boldsymbol{e}$, and $\delta_{\boldsymbol{e}} P_t$ is the distribution of X_t.

Denote by α_0 the best rate in $\|\delta_{\boldsymbol{e}} P_t - \pi\|_{\mathrm{TV}}$ convergence. It is known that for ergodic birth and death processes $\mathrm{gap}(\boldsymbol{Q}) = \alpha_0$; see, e.g. [46] or [8, Theorem 5.3]. We will point out (Section 3.1) that we have this equality also for ergodic reversible (unreliable) Jackson networks.

It is usually a very difficult (if not impossible) task to compute $\mathrm{gap}(\boldsymbol{Q})$. Sometimes it is possible to prove that $\mathrm{gap}(\boldsymbol{Q}) > 0$) (the existence) without being able to give computable bounds on the gap. We consider the problem of finding computable bounds for the L_2 spectral gap of unreliable Jackson network Markov processes which we will define later by the corresponding generators.

There is extensive literature on the speed of convergence to stationarity for general processes $\boldsymbol{X}$. Let us recall a few references. In order to prove the existence of the spectral gap for $\boldsymbol{X}$ it is possible to use the theory of Harris recurrent Markov processes, utilizing Lyapunov functions with appropriate drift conditions; see Meyn and Tweedie [40]. However, computable bounds are not easily obtainable by the Harris recurrence techniques. Some exceptions are known such as, for example, when $\mathbb{E} = \mathbb{R}$ (totally ordered state space) and in addition when the process is stochastically monotone; see [38] and [41]. Other approaches are possible via coupling methods or renewal theory methods; see, e.g. [2], [3], [4], and [10]. Sharper results leading to bounds on the spectral gap are possible via strong stationary times, strong stationary duality, Cheeger-type inequalities, Poincare inequalities, or direct spectral representations for the semi-group $(P_t,\ t > 0)$; see, e.g. [16], [17], [19], [22], [23], [33]–[35], and [37], and the book [10]. Symmetry assumptions turned out to be especially effective in analysis, and the reversibility of $\boldsymbol{X}$ is a typical assumption for many results. However, even for birth and death processes analysis of spectra and transient behaviour of $(P_t,\ t > 0)$ is far from being simple; see, e.g. [6], [7], [11], [26], [31], [32], [36], [45], [48], and [49], for some results on bounds on the gap, and [18], [25], and [37] for strong stationary times and duals approach to finite-state birth and death processes.

Jackson network processes can be seen as a generalization of birth and death processes, and one can expect that bounds for the spectral gap of a network should be related to some bounds on spectral gaps for some related birth and death processes. In fact, Jackson network processes are much more complicated than birth and death processes because they are built upon an additional Markov chain which guides the routeing inside the network. Reversibility for Jackson networks depends upon reversibility of the routeing matrix. It is known that the simplest Jackson networks with constant service rates are stochastically monotone (under coordinate-wise ordering), but in general the stochastic monotonicity depends on the properties of the corresponding state dependent service rates; see, e.g. [12] for many monotonicity properties of Jackson networks. Unfortunately, for unreliable Jackson networks no reasonable stochastic monotonicity is present (see, e.g. [15]), therefore known methods to find computable bounds on the spectral gap, using the stochastic monotonicity property, are not applicable for networks (also because all known results on computable bounds with a use of stochastic monotonicity require totally ordered state spaces). A plausible expectation is that the speed of convergence to stationarity of a network should correspond to a bottleneck node of the network. Some partial results in this direction can be found for networks with state independent service rates in [1] (for finite capacity networks), [5], and [20] (for tandems). For networks with state independent rates Lyapunov drift functions were studied in [21] and [27].

A related line of research is to study the essential spectrum of the generator $\boldsymbol{Q}$ of $(P_t,\ t > 0)$. A broad view on this topic can be found in [51]. The generator $\boldsymbol{Q}$ can act as operator on various function spaces (Banach lattices), such as, for example, $L^p,\ p \geq 1$, and the corresponding essential spectral gap is always larger than the gap defined by the underlying norm in a given function space. The essential spectral radius is directly related to large deviation principle (LDP) theory, to Lyapunov functions, and to compact sets with asymptotic results for the tail distributions of (the first) returning times. Finding the essential spectral radius for the L^2 space gives at once an upper bound on the speed of convergence in L^2, which is interesting, but more interesting for assessing the speed of convergence is to have lower bounds on the gap. In general, we do not know the results that characterize when the L^2 spectral radius is equal to the corresponding essential spectral radius, however some examples showing this equality for some ergodic birth and death processes are known; see, e.g. [51, Example 8.4]. For ergodic birth and death processes with constant (state independent) rates the L^2 spectral gap is known; see, e.g. examples after [9, Corollary 1.3]. For ergodic birth and death processes with constant rates the essential L^2 spectral gap is also known; see, e.g. in the context of Jackson networks, [28] and [29]. In the language of queueing processes, for an ergodic single M/M/1 station the L^2 spectral gap and the L^2 essential spectral gap are both equal to $(\sqrt{\lambda} - \sqrt{\mu})^2$. It would be interesting to characterize the class of networks for which this equality holds true more generally. The fact that the problem of using spectral theory to characterize rates of convergence is a rather complex problem, even for countable Markov chains used in queueing theory, can be seen, for example, from [10], [39], or [50].

Positive lower bounds for the spectral gap of Jackson networks with state dependent service rates were obtained via some related birth and death processes in [30] by using conductance bounds from [34]. A related comparison result for spectral gaps for classical Jackson networks is given in [14, Proposition 3.6], where a direct comparison involving the spectral gaps for some related birth and death processes is given using an additional assumption on the routeing. In this paper we give some bounds on the spectral gap for networks with state dependent service rates using Cheeger-type constants from the approach of [34], similarly as in [30], but related to some other birth and death processes than those defined in [30]. We consider in addition the possibility of having unreliable nodes. Unreliable Jackson networks are networks where in some subsets of the set of nodes the service stations can be broken and then repaired during the time evolution of the system. The breakdown and repair events can be of a rather general nature, but driven by a Markov process. In the time intervals when nodes are broken, there are several rules for rerouteing. For full details of such networks see [44] and [43]. We assume the property of reversibility for unreliable networks, however this assumption can be skipped if the nodes are reliable. In a few examples we compare our bounds with bounds obtainable from the results of [14] (lower bounds) and [29] (upper bounds). Jackson networks possess two remarkable properties crucial for our analysis, namely, the stationary distribution has a product form (also for unreliable networks) and exponential ergodicity that is directly related to the strong light-tailness of the stationary distribution. It is worth mentioning that admitting service rates which are state dependent in the model implies that each discrete distribution with the support $\{0, 1, 2, \ldots\}$ can appear as the stationary distribution for a node in the network. We will characterize light-tailness of the stationary distribution by the corresponding discrete hazard rate functions. The stationary distribution can be also characterized by the corresponding so-called equilibrium rates which turn out to be equal to individual, state dependent traffic intensity functions for the nodes of a network. Roughly speaking, the speed of convergence for a network will depend on a joint effect of how heavy the tails of the marginals of the stationary distribution are, together with how fast each single node operates, which in turn depends on the routeing in the network.

The paper is organized as follows. In the next section we introduce unreliable networks by giving the respective generator. In Section 3 we present a result relating the existence of the spectral gap of unreliable networks with the tail properties of its stationary distribution. In Section 4 we use equilibrium rates to reformulate our results from Section 3. In Section 5 we provide the proofs of the results from Section 3. Finally, in Section 6 we provide some examples of bounds on the spectral gap for networks.

2. Description of the network process

The classical *Jackson network* consists of m numbered servers, denoted by $M := \{1, \ldots, m\}$. Station $j \in M$ is a single server queue with infinite waiting room under the FCFS (first come first served) discipline. All the customers in the network are indistinguishable. There is an external Poisson arrival stream with intensity λ and arriving customers are sent to node j with probability r_{0j}, $\sum_{j=1}^{m} r_{0j} = r \le 1$. Customers arriving at node j from the outside or from other nodes request a service which is at node j provided with intensity $\mu_j(n)$ ($\mu_j(0) := 0$), where n is the number of customers at node j including the one being served. All service times and arrival processes are assumed to be independent.

A customer departing from node i immediately proceeds to node j with probability $r_{ij} \ge 0$ or departs from the network with probability r_{i0}. The routeing is independent of the past of the system, given the momentary node where the customer is. We assume that the stochastic matrix $R := (r_{ij},\ i, j \in M \cup \{0\})$ is irreducible.

Let $Z_j(t)$ be the number of customers present at node j at time $t \ge 0$. Then

$$Z(t) = (Z_1(t), \ldots, Z_m(t))$$

is the joint queue length vector at time instant $t \ge 0$ and $\boldsymbol{Z} := (Z(t),\ t \ge 0)$ is the joint queue length process with the state space $\mathbb{E} = \mathbb{Z}_+^m$.

The unique stationary distribution for $\boldsymbol{Z}$ exists if and only if the unique solution of the *traffic equation*

$$\lambda_i = \lambda r_{0i} + \sum_{j=1}^{m} \lambda_j r_{ji}, \qquad i = 1, \ldots, m \tag{2.1}$$

satisfies

$$C_i := 1 + \sum_{n=1}^{\infty} \frac{\lambda_i^n}{\prod_{y=1}^{n} \mu_i(y)} < \infty, \qquad 1 \le i \le m.$$

The parameters of a Jackson network are the arrival intensity λ, the routeing matrix R (with the corresponding traffic arrival intensities vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$), the vector of service rates $\boldsymbol{\mu} = (\mu_1(\cdot), \ldots, \mu_m(\cdot))$, and the number of servers m. Our standing assumption for all considered networks is that for all j, $\underline{\mu}_j := \inf_{n\geq 1} \mu_j(n) > 0$. We denote the overall minimal service intensity by $\underline{\mu} = \min_j \underline{\mu}_j$.

Assume now that the servers at the nodes in the Jackson network are unreliable, i.e. the nodes may break down. The breakdown event may occur in different ways. Nodes may break down as an isolated event or in groups simultaneously, and the repair of the nodes may end for each node individually or in groups as well. It is not required that those nodes which stopped service simultaneously return to service at the same time instant. To describe the system's evolution we have to enlarge the state space for the network process as we will describe below. Denote by $M_0 := \{0, 1, \dots, m\}$ the set of nodes enlarged by adding the *outside* node.

Let $D \subseteq M$ be the set of servers out of order, i.e. in *down status*.

- If $I \subseteq M \setminus D$, $I \neq \varnothing$ is a subset of nodes in *up status* then all servers in I break down simultaneously with intensity $\alpha^D_{D\cup I}(n_i : i \in M)$.
- If $H \subseteq D$, $H \neq \varnothing$ then all servers from H return from repair simultaneously with intensity $\beta^D_{D\setminus H}(n_i : i \in M)$.
- The routeing is changed according to the so-called *repetitive service-random destination blocking* (RS-RD blocking) rule: for the D-set of servers under repair routeing probabilities are restricted to nodes from $M_0 \setminus D$ as follows:

$$r^D_{ij} = \begin{cases} r_{ij}, & i, j \in M_0 \setminus D,\ i \neq j, \\ r_{ii} + \sum_{k\in D} r_{ik}, & i \in M_0 \setminus D,\ i = j. \end{cases}$$

  The external arrival rates are

$$\lambda r^D_{0j} = \lambda r_{0j} \quad \text{for nodes } j \in M \setminus D,$$

  and 0, otherwise.

Let $R^D = (r^D_{ij})_{i,j\in M_0\setminus D}$ be the modified routeing. Note that $R^{\varnothing} = R$.

We assume that for the intensities of breakdowns and repairs $\varnothing \neq I \subseteq M \setminus D$ and $\varnothing \neq H \subseteq D$ that

$$\alpha^D_{D\cup I}(n_i : i \in M) := \frac{\psi(D \cup I)}{\psi(D)},$$
$$\beta^D_{D\setminus H}(n_i : i \in M) := \frac{\phi(D)}{\phi(D \setminus H)},$$

where ψ and ϕ are arbitrary positive functions, defined for all subsets of the set of nodes, and $\psi(\varnothing) = \phi(\varnothing) = 1$. That means that breakdown and repair intensities depend on the sets of servers but are independent of the particular numbers of customers present in these servers.

In order to describe unreliable Jackson networks we need to attach to the state space $\mathbb{Z}^m_+$ of the corresponding standard network process an additional component which includes information on the availability of the system. We consider a new state space

$$\boldsymbol{n} = (D, n_1, n_2, \dots, n_m) \in \mathcal{P}(M) \times \mathbb{Z}^m_+ =: \mathbb{E},$$

where $\mathcal{P}(M)$ denotes the powerset of M. The first (0) coordinate in $\boldsymbol{n}$ we call the availability coordinate.

The set D is the set of servers in *down status*. At node $i \in D$ there are n_i customers waiting for server to be repaired. Denote possible transitions by

$$\begin{aligned}
T_{ij}\boldsymbol{n} &:= (D, n_1, \dots, n_i - 1, \dots, n_j + 1, \dots, n_m),\\
T_{0j}\boldsymbol{n} &:= (D, n_1, \dots, n_j + 1, \dots, n_m),\\
T_{i0}\boldsymbol{n} &:= (D, n_1, \dots, n_i - 1, \dots, n_m),\\
T_H\boldsymbol{n} &:= (D \setminus H, n_1, \dots, n_m),\\
T^I\boldsymbol{n} &:= (D \cup I, n_1, \dots, n_m).
\end{aligned}$$

Definition 2.1. The Markov process $\boldsymbol{X} = (\boldsymbol{X}(t),\ t \geq 0)$ defined by the infinitesimal generator

$$\begin{aligned}
\boldsymbol{Q}f(\boldsymbol{n}) &= \sum_{j=1}^{m}[f(T_{0j}\boldsymbol{n}) - f(\boldsymbol{n})]\lambda r_{0j}^D + \sum_{i=1}^{m}\sum_{j=1}^{m}[f(T_{ij}\boldsymbol{n}) - f(\boldsymbol{n})]\mu_i(n_i)r_{ij}^D\\
&\quad + \sum_{\varnothing \neq I \subseteq M\setminus D}[f(T^I\boldsymbol{n}) - f(\boldsymbol{n})]\frac{\psi(D \cup I)}{\psi(D)} + \sum_{\varnothing \neq H \subseteq D}[f(T_H\boldsymbol{n}) - f(\boldsymbol{n})]\frac{\phi(D)}{\phi(D \setminus H)}\\
&\quad + \sum_{j=1}^{m}[f(T_{j0}\boldsymbol{n}) - f(\boldsymbol{n})]\mu_j(n_j)r_{j0}^D
\end{aligned} \tag{2.2}$$

is called an *unreliable Jackson network*.

We denote the corresponding transition intensities (written in matrix form) by $[q(\boldsymbol{n},\boldsymbol{n}')]_{\boldsymbol{n},\boldsymbol{n}'\in\mathbb{E}}$.

Similarly to the classical case, the invariant distribution for this Markov process can be written in a product form.

Theorem 2.1. (See Sauer and Daduna [44].) *Let $\boldsymbol{X}$ be an unreliable Jackson network following the RS-RD blocking. If the routeing matrix R is reversible, i.e.*

$$\lambda_j r_{ji} = \lambda_i r_{ij}, \qquad i, j \in M$$

then the stationary distribution of process $\boldsymbol{X}$ is given by

$$\pi(\boldsymbol{n}) = \pi(D, n_1, \dots, n_m) = \frac{1}{C}\frac{\psi(D)}{\phi(D)}\prod_{i=1}^{m}\pi_i(n_i), \tag{2.3}$$

where

$$\pi_i(n_i) = \frac{1}{C_i}\frac{\lambda_i^{n_i}}{\prod_{k=1}^{n_i}\mu_i(k)}, \qquad C_i = 1 + \sum_{n=1}^{\infty}\frac{\lambda_i^n}{\prod_{y=1}^{n}\mu_i(y)} \tag{2.4}$$

and C is the normalization constant used for the availability coordinate. Constants C_i, $i = 1, \dots, m$ are all finite if and only if the network is ergodic.

Note that in this generality, the reduced state vector associated to the number of customers alone, without the availability coordinate, does not form a Markov process. The model of an unreliable network is an analogue of the classical Jackson network model but it can not be reduced to the classical one by adjusting parameters of the availability coordinate since all configurations of down nodes are possible with positive probability under our assumptions.

2.1. Equilibrium rate and hazard rate for a stationary distribution

For a nonnegative random variable $X \in \mathbb{Z}_+$, with probability function $p(k) = \mathbb{P}\{X = k\}$, such that for any $k \in \mathbb{Z}_+$, $\mathbb{P}\{X = k\} > 0$, the *total hazard function* H_p is defined for all $x \geq 0$ by

$$H_p(x) = -\log \bar{F}(x).$$

Further, the *discrete hazard function* we define for natural arguments by

$$h_p(k) = \frac{p(k)}{\bar{F}(k-1)}, \qquad k \geq 0,$$

where $\bar{F}(k) = \mathbb{P}\{X > k\}$. Note that for such a variable, for natural arguments $k \geq 0$,

$$H_p(k) = -\log \prod_{j=0}^{k} (1 - h_p(j)),$$

and for arbitrary $x \geq 0$ we have

$$H_p(x) = -\log \prod_{j=0}^{\lfloor x \rfloor} (1 - h_p(j)) = \sum_{j=0}^{\lfloor x \rfloor} \log\left(\frac{1}{1 - h_p(j)}\right), \tag{2.5}$$

where $\lfloor x \rfloor$ denotes the integer part of x and h_p is the hazard function.

Definition 2.2. We say that a discrete distribution $(p(k),\ k = 0, 1, \ldots)$ (or a discrete random variable X) is *strongly light-tailed* if there exists $\varepsilon > 0$ such that $\inf_{k \geq 0} h_p(k) > \varepsilon$.

The following lemma and example explain how *the strong light-tailness* and *the usual light-tailness* are related. Recall the usual light-tailness. An arbitrary distribution function F with its support contained in $[0, \infty)$ is light-tailed if $\int_0^\infty \mathrm{e}^{sx}\, \mathrm{d}F(x) < \infty$ for some $s > 0$.

Lemma 2.1. *Consider a random variable $X \in \mathbb{Z}_+$, with probability function $p(k) = \mathbb{P}\{X = k\}$, such that for any $k \in \mathbb{Z}_+$, $\mathbb{P}\{X = k\} > 0$, and p is strongly light-tailed. Then it is light-tailed in the usual sense.*

Proof. It is known (see, e.g. [42, Theorem 2.3.1]) that

$$\liminf_{x \to \infty} -\frac{1}{x} \log(\bar{F}(x)) > 0$$

implies that F is light-tailed. Note that

$$\frac{H_p(x)}{x} \geq \frac{H_p(\lfloor x \rfloor)}{\lfloor x \rfloor + 1}$$

for all $x \geq 0$, therefore,

$$\inf_n \frac{H_p(n)}{n+1} > 0 \Longrightarrow \liminf_{x \to \infty} \frac{H_p(x)}{x} > 0. \tag{2.6}$$

From the exponential light-tailness we have for all j, $\log(1/(1 - h_p(j))) > \log(1/(1 - \varepsilon))$, and, hence, from (2.5),

$$\frac{H_p(n)}{n+1} > \log\left(\frac{1}{1 - \varepsilon}\right) > 0,$$

which from (2.6) implies that F is light-tailed.

We now provide a simple example in order to demonstrate that for discrete distributions, strong light-tailness is a strictly stronger notion than the usual light-tailness. (This example shows at the same time that there exists a birth and death process having its rate of convergence to stationarity not exponentially fast, but having its stationary distribution light-tailed).

Example 2.1. Let us take as p the distribution which corresponds to the hazard function h_p given by $h_p(1) = \frac{1}{2}$,

$$h_p(k) = \begin{cases} \dfrac{1}{k} & \text{if } k = 2n+1,\ n \geq 1, \\ \dfrac{1}{2} & \text{if } k = 2n,\ n \geq 0. \end{cases}$$

This distribution is not strongly light-tailed since $\inf_k h_p(k) = 0$. However, for each natural n, $\lim_{n\to\infty} H_p(2n+1)/(2n+2) = \lim_{n\to\infty} H_p(2n)/(2n+1) = \log(2)/2 > 0$, and from (2.6) we obtain that p is light-tailed.

For a nonnegative random variable $X \in \mathbb{Z}_+$, with probability function $p(k) = \mathbb{P}\{X = k\}$, such that for any $k \in \{0, 1, 2, \ldots\}$, $\mathbb{P}\{X = k\} > 0$, we define the *equilibrium rate* function for $k \in \mathbb{Z}_+$ by

$$e_p(k) = \begin{cases} \dfrac{p(k+1)}{p(k)} & \text{if } k \geq 0, \\ 0 & \text{if } k < 0. \end{cases}$$

Since under our assumptions the equilibrium rate function $(e_p(k),\ k \geq 0)$ uniquely determines the probability function $(p(k),\ k \geq 0)$, it is therefore possible to express strong light-tailness in terms of equilibrium rates. The following equations connect hazard and equilibrium rate functions:

$$e_p(k) = \frac{h_p(k+1)(1 - h_p(k))}{h_p(k)}, \qquad k \geq 0 \tag{2.7}$$

and

$$h_p(k) = \frac{1}{1 + \sum_{j=k}^{\infty} e_p(k) \cdots e_p(j)}, \qquad k \geq 0. \tag{2.8}$$

It is worth mentioning that each discrete distribution with the support $\mathbb{Z}_+$ can appear as the stationary distribution for a birth and death process with constant birth rates and variable death rates. Strong light-tailness of π_i can be expressed in terms of the corresponding equilibrium rates, which in turn are equal to the corresponding birth/death ratios. We provide a precise formulation for a single birth and death process in the following lemma.

Lemma 2.2. *Consider $\{p(k)\}_{k\geq 0}$, an arbitrary probability function on $\mathbb{Z}_+$, such that $p(k) > 0, k \geq 0$, with the corresponding equilibrium rate $e_p(k),\ k \geq 0$. Then for each birth and death process $\mathbf{Z}$ with fixed $\lambda(k) \equiv \lambda > 0,\ k \geq 0$, and death rates defined by*

$$\frac{\lambda}{\mu(k+1)} = e_p(k), \qquad k \geq 0,$$

the stationary distribution of $\mathbf{Z}$ is equal to $p(k),\ k \geq 0$.

Proof. For the stationary distribution $\check{\pi}$ of the birth death process $\mathbf{Z}$ we have

$$\frac{\check{\pi}(i)}{\check{\pi}(0)} = \frac{\lambda^i}{\mu(1)\cdots\mu(n)} = \lambda^i \left[\lambda^i \frac{p(0)}{p(1)} \frac{p(1)}{p(2)} \cdots \frac{p(i-1)}{p(i)}\right]^{-1} = \frac{p(i)}{p(0)}, \qquad i \geq 1.$$

Thus, we have $p = \check{\pi}$.

Neither $h_p(k)$ nor $e_p(k)$ have to be convergent as $k \to \infty$. However, from (2.7) and (2.8) we obtain a connection between these limits if they exist and are finite.

Lemma 2.3. *Consider $\{p(k)\}_{k\geq 0}$, an arbitrary probability function on $\mathbb{Z}_+$, such that $p(k) > 0, k \geq 0$, with the corresponding equilibrium rate $e_p(k),\ k \geq 0$. Then $h_p = \lim_{k\to\infty} h_p(k)$ exists and $h_p \in (0, 1)$ if and only if $e_p = \lim_{k\to\infty} e_p(k)$ exists and $e_p \in (0, 1)$. In this case*

$$h_p = 1 - e_p.$$

Example 2.2. Recall that the negative binomial distribution is defined by

$$p(k) = \binom{r+k-1}{k}(1-p)^k p^r, \qquad r > 0,\ k = 0, 1, \dots, p \in (0, 1).$$

The corresponding equilibrium rate is given by

$$e_p(k) = \frac{(1-p)(k+r)}{(k+1)}, \qquad k = 0, 1, \dots.$$

The corresponding limit at ∞ fulfills $e_p = (1-p)$, and for the corresponding limit at ∞ of the hazard rate we obtain $h_p = p > 0$, which means that this distribution is strongly light-tailed.

Example 2.3. For the Poisson distribution

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad \lambda > 0,\ k = 0, 1, \dots,$$

and

$$e_p(k) = \frac{\lambda}{k+1}.$$

For the corresponding limits at infinity we have here $e_p = 0$, and $h_p = 1$, the Poisson distribution is strongly light-tailed.

It is worth mentioning that the negative binomial and Poisson distributions fit into the so-called Panjer recurrence scheme; more precisely, we say that $p(k)$ fulfills Panjer's recurrence if, for some $a, b \in \mathbb{R}$,

$$p(k+1) = \left(a + \frac{b}{k+1}\right)p(k), \qquad k = 0, 1, \dots,$$

which is equivalent to saying that the corresponding equilibrium rate has a hyperbolic form

$$e_p(k) = a + \frac{b}{k+1}.$$

For the negative binomial distribution $a := 1 - p$ and $b := (r-1)(1-p)$. In both cases the equilibrium rate function is monotone. Distributions with nonincreasing equilibrium rates are equivalently called PF_2 densities, for more details in connection with queueing networks; see [13].

Example 2.4. A discrete analog of the Pareto distribution can be defined by

$$p(k) = C\frac{1}{(k+1)^\alpha}, \qquad \alpha > 1,\ k = 0, 1, \dots,$$

where C is the normalization constant. Then

$$e_p(k) = \left(\frac{k+1}{k+2}\right)^{\alpha}.$$

For the corresponding limits at ∞ we have here $e_p = 1$, and $h_p = 0$, this distribution is heavy-tailed.

In the context of unreliable queueing networks it is natural to define the ratio $\lambda_i/\mu_i(k+1)$, where k is a variable, as *the traffic intensity function* for the ith station. From Lemma 2.2 it follows that for ergodic networks the traffic intensity function at the ith station is equal to the equilibrium rate of the marginal distribution π_i of the network's stationary distribution π. If we assume that the service intensity at node i is nondecreasing as a function of the number of customers at this node then π_i has a PF_2 density, and it is strongly light-tailed. Another possibility is that the traffic intensity function is increasing to 1 at a selected node i, and the network is ergodic but having at the node i a heavy-tailed distribution π_i. In the next section we will show that in such a case the network process will not converge to stationarity geometrically fast. Also, if at a fixed station i the traffic intensity function is not monotone and corresponds to a light-tailed distribution which is not strongly light-tailed as in Example 2.1, then such a network also will not converge to stationarity geometrically fast.

3. Existence of a spectral gap and light-tailed distributions

Theorem 3.1. (i) *Let $\boldsymbol{X}$ be an ergodic unreliable Jackson network process following the RS-RD blocking, with the infinitesimal generator $\boldsymbol{Q}$. Suppose that $\boldsymbol{Q}$ is bounded and the minimal service intensity is $\underline{\mu} > 0$.*

If the routeing matrix R is reversible then $\text{gap}(\boldsymbol{Q}) > 0$ *if and only if all distributions $\pi_i,\ i = 1, \ldots, m$ are strongly light-tailed.*

(ii) *Let $\boldsymbol{Z}$ be an ergodic classical Jackson network process with the corresponding infinitesimal generator $\boldsymbol{Q}(\boldsymbol{Z})$. Suppose that $\boldsymbol{Q}(\boldsymbol{Z})$ is bounded and the minimal service intensity $\underline{\mu} > 0$.*

Then $\text{gap}(\boldsymbol{Q}(\boldsymbol{Z})) > 0$ *if and only if all distributions $\pi_i,\ i = 1, \ldots, m$ are strongly light-tailed.*

The proof of this theorem will be given in Section 5.

We formulated the results on the positivity of the spectral gap and on the convergence to stationarity in terms of the discrete hazard functions of the stationary distribution. For queueing networks it would be however more reasonable to formulate the assumptions in terms of the parameters of the network.

The existence of the spectral gap of an unreliable network can be formulated in terms of the corresponding arrival and service rates (as a consequence of Theorem 3.1 and Lemma 2.2) as follows.

Corollary 3.1. *Let $\boldsymbol{X}$ be an ergodic unreliable Jackson network process following the RS-RD blocking, with the infinitesimal generator $\boldsymbol{Q}$. Suppose that $\boldsymbol{Q}$ is bounded and the minimal service intensity $\underline{\mu} > 0$. If the routeing matrix R is reversible then* $\text{gap}(\boldsymbol{Q}) > 0$ *if and only if for each $i = 1, \ldots, m$,*

$$\inf_k \left[1 + \sum_{j=k+1}^{\infty} \frac{\lambda_i^{j-k}}{\mu_i(k+1)\cdots\mu_i(j)}\right]^{-1} > 0.$$

In particular for ergodic networks, if for all $i = 1, \dots, m$*, the limits for the traffic intensity functions* $\lim_{k\to\infty} \lambda_i/\mu_i(k) < 1$ *exist then* $\text{gap}(\boldsymbol{Q}) > 0$.

For the classical Jackson network the assumption on reversibility can be skipped.

3.1. Speed of convergence to stationarity

Denote by α_0 the best rate in $\|\delta_{\boldsymbol{e}} P_t - \pi\|_{\text{TV}}$ convergence. It is known that for ergodic birth and death processes $\text{gap}(\boldsymbol{Q}) = \alpha_0$; see, e.g. [46] or [8, Theorem 5.3]. From [10, Theorem 8.8, Equation (2)], for ergodic reversible processes it is known that $\alpha_0 \geq \text{gap}(Q)$. From [10, Theorem 8.13, Equation (4)], we have the following result.

Theorem 3.2. *Let* $\boldsymbol{X}$ *be an ergodic, unreliable Jackson network following the RS-RD blocking, with generator* $\boldsymbol{Q}$*, given by (2.2), and the corresponding transition semigroup* (P_t)*. Suppose the routeing matrix* R *is reversible.*

If π_i *is strongly light-tailed for each* $i = 1, \dots, m$ *then the following conditions are equivalent:*

(i) *for all* $f \in L^2(\mathbb{E}, \pi)$

$$\|P_t f - \pi(f)\| \leq \mathrm{e}^{-\text{gap}(\boldsymbol{Q})t} \|f - \pi(f)\|, \qquad t > 0,$$

(ii) *for each* $\boldsymbol{e} \in \mathbb{E}$ *there exists* $C(\boldsymbol{e}) > 0$ *such that*

$$\|\delta_{\boldsymbol{e}} P_t - \pi\|_{\text{TV}} \leq C(\boldsymbol{e})\mathrm{e}^{-\text{gap}(\boldsymbol{Q})t}, \qquad t > 0,$$

where $\|\cdot\|_{\text{TV}}$ *denotes the total variation norm.*

Proof. First note that the network process is reversible under the assumption that R is reversible. It is enough to check the assumptions of Theorem 8.13 and [10, Equation (4)]. Let

$$p_t(\boldsymbol{e}, \boldsymbol{e}') = \frac{d P_t(\boldsymbol{e}, \cdot)}{d\pi}(\boldsymbol{e}'), \qquad t > 0,\ \boldsymbol{e},\ \boldsymbol{e}' \in \mathbb{E}.$$

Then

$$p_{2s}(\boldsymbol{e}, \boldsymbol{e}) = \frac{\mathbb{P}\{\boldsymbol{X}(2s) = \boldsymbol{e} \mid \boldsymbol{X}(0) = \boldsymbol{e}\}}{\pi(\boldsymbol{e})}.$$

Hence, $p_{2s}(\cdot, \cdot) \in L^{(1/2)}_{\text{loc}}(\pi)$ (with the usual notation for $L^p(\pi)$ spaces as in [10]) if $\sum_{\boldsymbol{e}\in A\subset\mathbb{E}} (\pi(\boldsymbol{e}))^{(1/2)} < \infty$ for bounded A, which trivially holds. The set of bounded functions with compact support is (also trivially) dense in $L^2(\pi)$ since $\mathbb{E}$ is a discrete space.

Remark 3.1. For the classical Jackson networks, the reversibility assumption on the routeing matrix R can be relaxed in order to obtain the implication (i) $\Rightarrow$ (ii).

4. Bounds on the spectral gap

In this section we recall some bounds on the spectral gaps of birth and death processes. For a more complete description; see [10, Chapter 5], [11], [48], [49], and the references therein.

Let us recall [35, Theorem 3.7]. For convenience we give a simplified formulation of it to the case of state independent birth rates.

Theorem 4.1. (See Liggett [35].) *Assume that* $\mathbf{Z}$ *is an ergodic birth and death process on* $\mathbb{Z}_+$, *with state independent birth rates* $\lambda > 0$, *and possibly state dependent death rates* $\mu(n) > 0$, *and for all* $i \geq 0$, *and for some* $c, d > 0$, *we have*

$$\sum_{j>i} \pi(j) \leq c\pi(i)\lambda \quad \text{and} \quad \sum_{j>i} \pi(j) \leq d\pi(i).$$

Then for the corresponding generator $\boldsymbol{Q}(\mathbf{Z})$,

$$\operatorname{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \frac{(\sqrt{d+1} - \sqrt{d})^2}{c} \geq \frac{1}{2c(1+2d)}. \tag{4.1}$$

In the case of constant birth rates, from [35, Corollary 3.8], we have a necessary and sufficient condition for $\operatorname{gap}(\boldsymbol{Q}(\mathbf{Z}))$ to be positive is that the stationary distribution is such that

$$\inf_{i\geq 0} \frac{\pi(i)}{\sum_{j\geq i} \pi(j)} > 0,$$

which is by definition the strong light-tailness. Therefore from [35, Corollary 3.8] we have the following lemma.

Lemma 4.1. *Assume that* $\mathbf{Z}$ *is an ergodic birth and death process on* $\mathbb{Z}_+$, *with state independent birth rates* $\lambda > 0$, *and possibly state dependent death rates* $\mu(n) > 0$. *Then* $\operatorname{gap}(\boldsymbol{Q}(\mathbf{Z})) > 0$ *if and only if the stationary distribution* π *is strongly light-tailed. Moreover, if for some* $\varepsilon > 0$, *we have*

$$\inf_{n\geq 0} h_\pi(n) \geq \varepsilon$$

then

$$\operatorname{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \frac{\lambda(1 - \sqrt{1-\varepsilon})^2}{1-\varepsilon} \geq \frac{\lambda\varepsilon^2}{2(1-\varepsilon)(2-\varepsilon)}. \tag{4.2}$$

Proof. From $\sum_{j>i} \pi(j) \leq c\pi(i)\lambda$ we have $\sum_{j\geq i} \pi(j) \leq c\pi(i)\lambda + \pi(i)$, so for the lower bound on the hazard function we have $\varepsilon = 1/(1 + c\lambda)$, therefore $c = (1-\varepsilon)/(\lambda\varepsilon)$. Similarly, we obtain $d = (1-\varepsilon)/\varepsilon$, and using (4.1) we obtain (4.2).

A lower bound on the spectral gap can be given directly in terms of the birth and death rates; see, e.g. [48].

Lemma 4.2. *Assume that* $\mathbf{Z}$ *is an ergodic birth and death process on* $\mathbb{Z}_+$, *with state independent birth rates* $\lambda > 0$, *and possibly state dependent death rates* $\mu(n) > 0$. *Then*

$$\operatorname{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \inf_{n\geq 0} [\lambda + \mu(n+1) - \sqrt{\lambda\mu(n)} - \sqrt{\lambda\mu(n+1)}].$$

Remark 4.1. For more details on the estimation of spectral gaps for birth and death processes; see [9, Corollary 1.2 and Corollary 1.3] and also [10], [11], [47], [48], and [49]. It is natural to ask how do different bounds compare. It turns out that the optimality of a given bound strongly depends on the parameters of a given birth and death process, as described in an example after [8, Theorem (5.2)]. In a sense, different bounds are *incomparable*; as stated there. For particular cases it is reasonable to examine all possibilities.

Combining the above bounds for birth and death processes and the bounds obtained in the proof of Theorem 3.1 (see (5.9)) we have from (4.2) the following proposition.

Proposition 4.1. (i) *Let X be an ergodic, unreliable Jackson network following the RS-RD blocking, with generator $\boldsymbol{Q}$, given by (2.2). Suppose the routeing matrix R is reversible.*

If π_i is strongly light-tailed for each $i = 1, \dots, m$, and

$$\inf_{n \geq 0} h_{\pi_i}(n) \geq \varepsilon_i > 0,$$

then

$$\mathrm{gap}(\boldsymbol{Q}) \geq \frac{1}{8|\boldsymbol{Q}|}\left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\mathrm{gap}(\check{\boldsymbol{Q}}_0) \wedge \min_{1 \leq i \leq m} \lambda_i (1 - \sqrt{1 - \varepsilon_i})^2/(1 - \varepsilon_i)}{1 + \bar{d}\bar{b}(2m + 1)}\right)^2$$

and

$$\mathrm{gap}(\boldsymbol{Q}) \geq \frac{1}{8|\boldsymbol{Q}|}\left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\mathrm{gap}(\check{\boldsymbol{Q}}_0) \wedge \min_{1 \leq i \leq m} \inf_{n \geq 0}[\lambda_i + \mu_i(n + 1) - \sqrt{\lambda_i \mu_i(n)} - \sqrt{\lambda_i \mu_i(n + 1)}]}{1 + \bar{d}\bar{b}(2m + 1)}\right)^2,$$

where $\bar{d}$, $\bar{b}$, $|\boldsymbol{Q}|$, $q^{\min}$, $\check{q}^{\max}$ are defined by (5.7), (5.6), (5.2), (5.4), and (5.5), respectively.

(ii) *Let $\mathbf{Z}$ be an ergodic classical Jackson network process with the corresponding infinitesimal generator $\boldsymbol{Q}(\mathbf{Z})$. Suppose that $\boldsymbol{Q}(\mathbf{Z})$ is bounded and the minimal service intensity is $\underline{\mu} > 0$. If π_i is strongly light-tailed, for each $i = 1, \dots, m$, and*

$$\inf_{n \geq 0} h_{\pi_i}(n) \geq \varepsilon_i > 0,$$

then

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \frac{1}{8|\boldsymbol{Q}(\mathbf{Z})|}\left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\min_{1 \leq i \leq m} \lambda_i (1 - \sqrt{1 - \varepsilon_i})^2/(1 - \varepsilon_i)}{1 + \bar{b}2m}\right)^2 \tag{4.3}$$

and

$$\mathrm{gap}(\boldsymbol{Q}) \geq \frac{1}{8|\boldsymbol{Q}(\mathbf{Z})|}\left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\min_{1 \leq i \leq m} \inf_{n \geq 0}[\lambda_i + \mu_i(n + 1) - \sqrt{\lambda_i \mu_i(n)} - \sqrt{\lambda_i \mu_i(n + 1)}]}{1 + \bar{b}2m}\right)^2.$$

In all the above given bounds the factor $1 + \bar{d}\,\bar{b}(2m + 1)$ can be reduced to 1 if in the network $r_{i0} > 0$ and $r_{0i} > 0$ for all $i = 1, \dots, m$. The bounds obtained in the above proposition are valid for a quite general class of networks but it is reasonable to search for alternative bounds and alternative methods under some additional structural assumptions. We recall two cases for classical Jackson networks, the first one with state dependent service rates but fulfilling a partial balance requirement for the roueting matrix (see [14, Proposition 4.4]), the second one for classical Jackson networks with state independent service rates (see [29]).

Proposition 4.2. *Let $\mathbf{Z}$ be an ergodic classical Jackson network process with the corresponding infinitesimal generator $\boldsymbol{Q}(\mathbf{Z})$. Suppose that $\boldsymbol{Q}(\mathbf{Z})$ is bounded and the minimal service intensity is $\underline{\mu} > 0$. Assume that the routeing matrix R has strict positive departure probabilities $r_{i0} > 0$ and that $\lambda r_{0i} > 0$ for $i = 1, \dots, m$.*

Assume further a partial balance condition

$$\lambda_j \sum_{i=1}^{m} r_{ji} = \sum_{i=1}^{m} \lambda_i r_{ij} \quad \text{for all } j = 1, \dots, m. \tag{4.4}$$

Then

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \min_{1\leq i\leq m} \mathrm{gap}(\tilde{Q}_i),$$

where, for $i = 1, \ldots, m$, $\tilde{Q}_i$ *denotes the generator of the birth and death process with the birth rate* λr_{0i} *and the state dependent death rate* $\mu_i(n_i)r_{i0}$.

Corollary 4.1. *Under the assumptions of Proposition 4.2, if, in addition,* π_i *is strongly light-tailed for each* $i = 1, \ldots, m$, *and*

$$\inf_{n\geq 0} h_{\pi_i}(n) \geq \varepsilon_i > 0,$$

then

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \min_{1\leq i\leq m} \frac{\lambda r_{0i}(1-\sqrt{1-\varepsilon_i})^2}{1-\varepsilon_i}$$

and

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \geq \min_{1\leq i\leq m} \inf_{n\geq 0} [\lambda r_{0i} + \mu_i(n+1)r_{i0} - \sqrt{\lambda r_{0i}\mu_i(n)r_{i0}} - \sqrt{\lambda r_{0i}\mu_i(n+1)r_{i0}}].$$

Now we recall from [29] some special cases of classical Jackson networks in order to present some (upper) bounds on the corresponding L^2 spectral gap. The results of [29] are related to the essential spectral gap. In Section 6 we will compare our lower bounds with the upper bounds presented below and we will obtain in some cases a nice approximation for the L^2 spectral gap. Because the essential L^2 spectral gap is larger then the L^2 spectral gap from [29, Corollary 3.4 and Proposition 3.6].

Proposition 4.3. *Let* $\mathbf{Z}$ *be an ergodic classical Jackson network process with the corresponding infinitesimal generator* $\boldsymbol{Q}(\mathbf{Z})$. *Assume that the service intensities are state independent.*

(i) *If the routeing is completely symmetrical, i.e.* $r_{ij} = p < 1/(m-1)$ *for all* $i \neq j$, $i, j = 1, \ldots, m$, *and for some* $i_0 \in \{1, \ldots, m\}$ *we have*

$$\min_{1\leq i\leq m} (\sqrt{\mu_i} - \sqrt{\lambda_i}) = \sqrt{\mu_{i_0}} - \sqrt{\lambda_{i_0}}$$

and

$$\min_{1\leq i\leq m} \left(\frac{\mu_i}{\sqrt{\mu_{i_0}}} - \frac{\lambda_i}{\sqrt{\lambda_{i_0}}} \right) = \sqrt{\mu_{i_0}} - \sqrt{\lambda_{i_0}},$$

then

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \leq \left(1 - \frac{(m-1)p^2}{1-(m-2)p}\right) \min_{1\leq i\leq m} (\sqrt{\mu_i} - \sqrt{\lambda_i})^2.$$

(ii) *If* $m = 3$, *and*

$$R = \begin{pmatrix} 0 & r_{01} & r_{02} & r_{03} \\ 1-(p+q) & 0 & p & q \\ 1-(p+q) & q & 0 & p \\ 1-(p+q) & p & q & 0 \end{pmatrix}, \tag{4.5}$$

where $p, q \in (0, 1)$, $p + q < 1$, *then*

$$\mathrm{gap}(\boldsymbol{Q}(\mathbf{Z})) \leq \frac{1-p^3-q^3-3pq}{1-pq} \min_{1\leq i\leq m} (\sqrt{\mu_i} - \sqrt{\lambda_i})^2,$$

provided that $\lambda_i/\mu_i = \lambda_j/\mu_j$, $i, j \in M$, *or there exists* i_0 *such that* $\mu_i \geq \mu_{i_0}$ *and* $\lambda_i \leq \lambda_{i_0}$ *for all* i.

5. Proof of Theorem 3.1

We present the proof of Theorem 3.1 using the following theorem.

Theorem 5.1. (See [35, Theorem 2.6].) *Suppose that a pure jump Markov process $\boldsymbol{X}$, with generator $\check{\boldsymbol{Q}}$ and stationary distribution π evolves on the product state space $\mathbb{E} = \mathbb{E}_0 \times \mathbb{E}_1 \times \cdots \mathbb{E}_m$, $m \geq 1$, having coordinates which are independent Markov processes such that the ith coordinate has generator $\check{\boldsymbol{Q}}_i$, denumerable state space $\mathbb{E}_i$, and invariant probability measure π_i. Then π is the product measure of π_is and*

$$\operatorname{gap}(\check{\boldsymbol{Q}}) = \min_{0 \leq i \leq m} \operatorname{gap}(\check{\boldsymbol{Q}}_i).$$

Proof of Theorem 3.1(i). We assume that the availability coordinate process is not degenerate with ϕ and ψ positive. Let $\check{\boldsymbol{Q}}$ be the generator associated with an $(m+1)$-dimensional process $(\boldsymbol{Y}_t, \check{\boldsymbol{Z}}_t)_{t \geq 0}$, where $\check{\boldsymbol{Z}}_t$ is the vector of m independent birth and death processes with generators $\check{\boldsymbol{Q}}_i$, $i = 1, \ldots, m$, given by

$$\check{\boldsymbol{Q}}_i f(n) = [f(n+1) - f(n)]\lambda_i + [f(n) - f(n-1)]\mu_i(n), \qquad n \in \mathbb{N},$$

and $\boldsymbol{Y}_t$ is the process on state space $\mathcal{P}(M)$ with infinitesimal generator denoted by $\check{\boldsymbol{Q}}_0$ and the stationary distribution:

$$\pi_0(I) = \frac{1}{C} \frac{\psi(I)}{\phi(I)}, \qquad C := \left(\sum_{I \subseteq M} \frac{\psi(I)}{\phi(I)} \right).$$

We write $[\check{q}(\boldsymbol{n}, \boldsymbol{n}')]_{\boldsymbol{n}, \boldsymbol{n}' \in \mathbb{E}}$ for the corresponding transition intensities.

The stationary distribution of the process with generator $\check{\boldsymbol{Q}}_i$ is π_i, which is given in the product equation (2.4) for networks.

Consider the following Cheeger's constants for $A \subset \mathbb{E}$:

$$\kappa(A) := \frac{\sum_{\boldsymbol{n} \in A} \pi(\boldsymbol{n}) q(\boldsymbol{n}, A^c)}{\pi(A)\pi(A^c)}, \qquad \kappa := \inf_{A : \pi(A) \in (0,1)} \kappa(A),$$
$$\check{\kappa}(A) := \frac{\sum_{\boldsymbol{n} \in A} \pi(\boldsymbol{n}) \check{q}(\boldsymbol{n}, A^c)}{\pi(A)\pi(A^c)}, \qquad \check{\kappa} := \inf_{A : \pi(A) \in (0,1)} \check{\kappa}(A),$$

where π is given by (2.3).

We will show that there exist $0 < v_1, v_2 < \infty$ such that, uniformly for all $A \subset \mathbb{E}$,

$$v_2 \sum_{\boldsymbol{n} \in A} \pi(\boldsymbol{n}) \check{q}(\boldsymbol{n}, A^c) \geq \sum_{\boldsymbol{n} \in A} \pi(\boldsymbol{n}) q(\boldsymbol{n}, A^c) \geq v_1 \sum_{\boldsymbol{n} \in A} \pi(\boldsymbol{n}) \check{q}(\boldsymbol{n}, A^c). \tag{5.1}$$

Then with $0 < v_1, v_2 < \infty$ as in (5.1), we use [34, Theorem 2.1], and since the process with the generator $\check{\boldsymbol{Q}}$ is reversible, we have that $\operatorname{gap}(\check{\boldsymbol{Q}}) \leq \check{\kappa}$. Furthermore, uniformly in A, $\check{\kappa}(A) \leq (v_1)^{-1} \kappa(A)$, hence, $\check{\kappa} \leq (v_1)^{-1} \kappa$. Under our assumptions we will have $\operatorname{gap}(\check{\boldsymbol{Q}}) > 0$ which in turn, using [34, Theorem 2.3] (which assures that $\kappa^2/(8|\boldsymbol{Q}|) \leq \operatorname{gap}(\boldsymbol{Q})$) will imply that $\operatorname{gap}(\boldsymbol{Q}) > 0$. Here

$$|\boldsymbol{Q}| = \pi - \operatorname{ess\,sup}_{\boldsymbol{n}} q(\boldsymbol{n}, \{\boldsymbol{n}\}^c). \tag{5.2}$$

Similarly, it is possible to argue that $\operatorname{gap}(\boldsymbol{Q}) > 0$ implies that $\operatorname{gap}(\check{\boldsymbol{Q}}) > 0$.

In order to complete the proof we turn now to show the validity of (5.1), which is equivalent to

$$\inf_{\substack{A\subset\mathbb{E}\\ \pi(A)\in(0,1)}}\left\{\frac{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)}\right\}\geq v_1>0 \tag{5.3}$$

and

$$\sup_{\substack{A\subset\mathbb{E}\\ \pi(A)\in(0,1)}}\left\{\frac{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)}\right\}\leq v_2<\infty.$$

For a fixed A such that $\pi(A)\in(0,1)$, we define

$$\partial A=\{\boldsymbol{n}\in A\colon q(\boldsymbol{n},A^c)>0\},\qquad \partial\check{A}=\{\boldsymbol{n}\in A\colon \check{q}(\boldsymbol{n},A^c)>0\}.$$

Let

$$q^{\min}=\inf_{A:\pi(A)\in(0,1)}\inf_{\boldsymbol{n}\in\partial A}\{q(\boldsymbol{n},A^c)\},\qquad q^{\max}=\sup_{A:\pi(A)\in(0,1)}\sup_{\boldsymbol{n}\in\partial A}\{q(\boldsymbol{n},A^c)\}. \tag{5.4}$$

From our assumptions it follows that the generators are bounded and $\underline{\mu}>0$, therefore, $q^{\min}>0$ and $q^{\max}<\infty$.

For

$$\check{q}^{\min}=\inf_{A:\pi(A)\in(0,1)}\inf_{\boldsymbol{n}\in\partial\check{A}}\{\check{q}(\boldsymbol{n},A^c)\},\qquad \check{q}^{\max}=\sup_{A:\pi(A)\in(0,1)}\sup_{\boldsymbol{n}\in\partial\check{A}}\{\check{q}(\boldsymbol{n},A^c)\}, \tag{5.5}$$

we also have $\check{q}^{\min}>0$ and $\check{q}^{\max}<\infty$.

For each A such that $\pi(A)\in(0,1)$, we have

$$\frac{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in A}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)}=\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)},$$

so we obtain

$$\frac{q^{\max}}{\check{q}^{\min}}\cdot\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})}\geq\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)}\geq\frac{q^{\min}}{\check{q}^{\max}}\cdot\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})}.$$

We will continue our argument in the case of the lower bound (5.3). The existence of this lower bound ensures that if $\operatorname{gap}(\check{\boldsymbol{Q}})>0$ then $\operatorname{gap}(\boldsymbol{Q})>0$. Note that from Theorem 5.1 and Lemma 4.1, the inequality $\operatorname{gap}(\check{\boldsymbol{Q}})>0$ is equivalent to the condition that π_i is strongly light-tailed, for each $i=1,\dots,m$. The proof for the upper bound is similar and we skip it. In order to show (5.3) it is enough to check that

$$0<\inf_{A:\pi(A)\in(0,1)}\zeta(A),\quad \text{where } \zeta(A):=\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})}.$$

If the network is such that for all $i=1,\dots,m$, $r_{0i}>0$ and $r_{i0}>0$ then $\partial\check{A}\subseteq\partial A$. In that case $\inf_{A\colon\pi(A)\in(0,1)}\zeta(A)\geq 1$, and we can take $v_1=q^{\min}/\check{q}^{\max}$. Otherwise, we have to analyse $\partial\check{A}$, and ∂A in more detail.

Let us examine the difference between $\pi(\boldsymbol{n})$ and $\pi(\boldsymbol{n}')$ when $\boldsymbol{n}'$ and $\boldsymbol{n}$ differ exactly on one nonavailability coordinate by at most 1 and when $\boldsymbol{n}$ and $\boldsymbol{n}'$ have two different sets of broken nodes D and D', respectively.

Recall from (2.3) that for $\boldsymbol{n} = (D, n_1, \ldots, n_m) \in \mathscr{P}(M) \times \mathbb{Z}_+^m$ we have

$$\pi(\boldsymbol{n}) = \pi(D, n_1, \ldots, n_m) = \frac{1}{C}\frac{\psi(D)}{\phi(D)}\prod_{i=1}^{m}\pi_i(n_i), \quad \text{where } \pi_i(n_i) := \frac{1}{C_i}\frac{\lambda_i^{n_i}}{\prod_{y=1}^{n_i}\mu_i(y)}.$$

For $n_i \geq 1$,

$$\pi_i(n_i + 1) = \frac{1}{C_i}\frac{\lambda_i^{n_i+1}}{\prod_{y=1}^{n_i+1}\mu_i(y)} = \pi_i(n_i)\frac{\lambda_i}{\mu_i(n_i + 1)}$$

and

$$\pi_i(n_i - 1) = \frac{1}{C_i}\frac{\lambda_i^{n_i-1}}{\prod_{y=1}^{n_i-1}\mu_i(y)} = \pi_i(n_i)\frac{\mu_i(n_i)}{\lambda_i},$$

thus, using $\underline{\mu}_i := \inf_n \mu_i(n) > 0$ and $\bar{\mu}_i := \sup_n \mu_i(n) < \infty$, we have bounds

$$\frac{\lambda_i}{\bar{\mu}_i}\pi_i(n_i) \leq \pi_i(n_i + 1) \leq \pi_i(n_i)\frac{\lambda_i}{\underline{\mu}_i}, \qquad \frac{\underline{\mu}_i}{\lambda_i}\pi_i(n_i) \leq \pi_i(n_i - 1) \leq \pi_i(n_i)\frac{\bar{\mu}_i}{\lambda_i}.$$

Define

$$\bar{b} = \max_{1\leq i\leq m}\left(\frac{\bar{\mu}_i}{\lambda_i}\right), \qquad \underline{b} = \min_{1\leq i\leq m}\left(\frac{\lambda_i}{\bar{\mu}_i}\right), \tag{5.6}$$

$$\bar{d} = \max_{D_1\neq D_2}\frac{\psi(D_2)\phi(D_1)}{\phi(D_2)\psi(D_1)} \quad \text{and} \quad \underline{d} = \min_{D_1\neq D_2}\frac{\psi(D_2)\phi(D_1)}{\phi(D_2)\psi(D_1)}. \tag{5.7}$$

Then, if $\boldsymbol{n}$ and $\boldsymbol{n}'$ differ by at most 1 on exactly one coordinate $i \in \{1, \ldots, m\}$, and have sets D and D', respectively, on the availability coordinate, then

$$\underline{b}\pi_i(n_i) \leq \pi_i(n_i') \leq \bar{b}\pi_i(n_i)$$

and

$$\underline{d}\underline{b}\pi(\boldsymbol{n}) \leq \pi(\boldsymbol{n}') \leq \bar{d}\bar{b}\pi(\boldsymbol{n}). \tag{5.8}$$

We rewrite $\zeta(A)$ as

$$\zeta(A) = \frac{\sum_{\boldsymbol{n}\in\partial A\cap\partial\check{A}}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial\check{A}\setminus\partial A}\pi(\boldsymbol{n})}.$$

Let us consider $\boldsymbol{n} \in \partial\check{A} \setminus \partial A$. Then there exists some $\boldsymbol{n}' \in A^c$ such that original process with the intensity q cannot move there in one step, but the process with $\check{q}$ can. The state $\boldsymbol{n}'$ must be of the form $\boldsymbol{n}' = T_{0i_0}\boldsymbol{n}$ or $\boldsymbol{n}' = T_{j_0 0}\boldsymbol{n}$ (arrival or departure) since changing the availability coordinate is always possible in both processes, i.e. either both processes would leave A or none. We will analyse the case of arrival since in the case of departure we can argue analogously. The key observation in this argument is the following: if $\boldsymbol{n}' = T_{0i_0}\boldsymbol{n}$, but the arrival intensity to node i_0 is equal to 0 for the network process or this arrival movement is blocked by D, then the node i_0 must be reachable by an *unblocking* movement $D \to \varnothing$ and then a T_{0i_0} transition, or by an unblocking movement $D \to \varnothing$ and then an arrival to some station different from i_0, and a migration movement or a series of consecutive migration movements. There are possibly multiple paths, but we can search for the minimal ones (which can be multiple with the same length). Intuitively speaking, we search for the shortest connection to a source node (i.e. a node which admits arrivals from the outside) from the i_0 node (in the case of departure movement $\boldsymbol{n}' = T_{j_0 0}\boldsymbol{n}$ we search for the shortest connection to a sink node). Consider all shortest paths of movements that connect $\boldsymbol{n}$ with $\boldsymbol{n}'$ in the network. Denote such a path by $\boldsymbol{n} = \boldsymbol{n}_0, \boldsymbol{n}_1 = T_D\boldsymbol{n}_0, \dots, \boldsymbol{n}_k = T^D\boldsymbol{n}_{k-1} = \boldsymbol{n}'$ for $k \le m+1$. Note that each such path is not greater than $m+1$ since we can take as the first transition the one which puts D to $\varnothing$ on the availability coordinate, and the worst case for the other transitions is when the station i_0 is the last station in a m-series network. Moreover, each state on the path differs from $\boldsymbol{n}$ by at most 1 on only one nonavailability coordinate (because on nonavailability coordinates an arrival changes one coordinate by $+1$, and consecutive transitions change coordinates in such a way that after a transition the resulting state has exactly one coordinate changed by $+1$). Furthermore, there exists a state $\boldsymbol{n}_j$ on this path such that the network process leaves A, and either $\boldsymbol{n}_j \in \partial\check{A} \cap \partial A$ or $\boldsymbol{n}_j \in \partial A \setminus \partial\check{A}$. Since $\boldsymbol{n}_j$ differs from $\boldsymbol{n}$ by at most 1 on exactly one coordinate, from (5.8) we have $\pi(\boldsymbol{n}) \le \bar{d}\bar{b}\pi(\boldsymbol{n}_j)$. If we take two points on the border $\partial\check{A} \setminus \partial A$ for which the coordinate-wise distance is large enough then the corresponding border points on ∂A defined above must be different because $\boldsymbol{n}_j$ always differs from $\boldsymbol{n}$ by at most 1 on a single coordinate. More precisely, let $\boldsymbol{n} \in \partial\check{A} \setminus \partial A$ and $\boldsymbol{m} \in \partial\check{A} \setminus \partial A$ be such that they are different by more than 2 on each coordinate, then the corresponding points $\boldsymbol{n}_j$ and $\boldsymbol{m}_{j'}$, elements of ∂A, are distinct. In order to give a very rough bound on $\sum_{\partial\check{A}\setminus\partial A}\pi(\boldsymbol{n})$ we observe that for a fixed $\boldsymbol{n}_j$ point there are not more than $2m+1$ points that are different by at most 1 on a single coordinate from $\boldsymbol{n}_j$, and $\boldsymbol{n}_j$ can potentially be on a transition (unblocking and migration) path described above for these points. Therefore, we have

$$\sum_{\partial\check{A}\setminus\partial A}\pi(\boldsymbol{n}) \le \bar{d}\bar{b}(2m+1)\bigg(\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n})\bigg)$$

and

$$\begin{aligned}\zeta(A) &\ge \frac{\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \bar{d}\bar{b}(2m+1)(\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n}))}\\
&\ge \frac{\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n})}{(1+\bar{d}\bar{b}(2m+1))(\sum_{\boldsymbol{n}\in\partial\check{A}\cap\partial A}\pi(\boldsymbol{n}) + \sum_{\boldsymbol{n}\in\partial A\setminus\partial\check{A}}\pi(\boldsymbol{n}))}\\
&= \frac{1}{1+\bar{d}\bar{b}(2m+1)}.\end{aligned}$$

Summing up, we obtain

$$\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})q(\boldsymbol{n},A^c)}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})\check{q}(\boldsymbol{n},A^c)} \ge \frac{q^{\min}}{\check{q}^{\max}}\frac{\sum_{\boldsymbol{n}\in\partial A}\pi(\boldsymbol{n})}{\sum_{\boldsymbol{n}\in\partial\check{A}}\pi(\boldsymbol{n})} \ge \frac{q^{\min}}{\check{q}^{\max}}\frac{1}{1+\bar{d}\bar{b}(2m+1)}$$

and

$$\check{\kappa}\frac{q^{\min}}{\check{q}^{\max}}\frac{1}{1+\bar{d}\bar{b}(2m+1)} \le \kappa,$$

which implies (using [34, Theorem 2.3])

$$\begin{aligned}\mathrm{gap}(\boldsymbol{Q}) &\ge \left(\check{\kappa}\frac{q^{\min}}{\check{q}^{\max}}\frac{1}{1+\bar{d}\bar{b}(2m+1)}\right)^2[(8|\boldsymbol{Q}|)]^{-1},\\
\mathrm{gap}(\boldsymbol{Q}) &\ge \left(\frac{q^{\min}}{\check{q}^{\max}}\frac{\mathrm{gap}(\check{\boldsymbol{Q}})}{1+\bar{d}\bar{b}(2m+1)}\right)^2[(8|\boldsymbol{Q}|)]^{-1}\end{aligned}$$

and, finally,

$$\operatorname{gap}(\boldsymbol{Q}) \geq \left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\min_{0\leq i\leq m} \operatorname{gap}(\check{\boldsymbol{Q}}_i)}{1+\bar{d}\bar{b}(2m+1)}\right)^2 [(8|\boldsymbol{Q}|)]^{-1}. \tag{5.9}$$

Proof of Theorem 3.1(ii). Note that we cannot specify the parameters of an ergodic unreliable Jackson network process $\boldsymbol{X}$ in order to obtain the classical ergodic Jackson network process $\boldsymbol{Z}$ as a special case. However, it is possible to repeat all steps in the proof of Theorem 3.1(i) for $\boldsymbol{Z}$ (skipping the availability coordinate, and reducing $2m+1$ to $2m$) to obtain

$$\operatorname{gap}(\boldsymbol{Q}(\boldsymbol{Z})) \geq \left(\frac{q^{\min}}{\check{q}^{\max}} \frac{\min_{1\leq i\leq m} \operatorname{gap}(\check{\boldsymbol{Q}}_i)}{1+\bar{b}2m}\right)^2 [(8|\boldsymbol{Q}(\boldsymbol{Z})|)]^{-1}.$$

6. Numerical examples

We will use two examples from [29] in order to estimate the L^2 spectral gap.

Example 6.1. Let $\boldsymbol{Z}$ be the classical Jackson network with $m = 3$ stations with the arrival intensity λ and the routeing matrix R given in (4.5) and with $r_{01} = r_{02} = r_{03} = \frac{1}{3}$, where $p, q \in (0,1)$, $p+q < 1$. Then $\lambda_1 = \lambda_2 = \lambda_3 = \lambda/(3(1-(p+q))$ is the solution to the traffic equation. Moreover, assume that service intensities are constant and are given by $\mu_i = c\lambda_i$, $i = 1,2,3$, where $c > 1$. The network is ergodic with a stationary distribution being the product of π_i, $i = 1,2,3$, where $\pi_i(k) = (1-1/c)(1/c)^k$, $i = 1,2,3$, $k = 0,1,\ldots$. The conditions of Proposition 4.3(ii) are fulfilled and we have:

$$\begin{aligned}
\operatorname{gap}(\boldsymbol{Q}(\boldsymbol{Z})) &\leq \operatorname{gap}_{\text{ess}} \\
&:= \frac{1-p^3-q^3-3pq}{1-pq}\lambda_1(\sqrt{c}-1)^2 \\
&= \frac{p^2+p-pq+q^2+q+1}{1-pq}\frac{\lambda}{3}(\sqrt{c}-1)^2,
\end{aligned}$$

where $\operatorname{gap}_{\text{ess}}$ denotes the essential spectral gap of the Wolf essential spectrum; see [51]. We will compare the above upper bound with the bounds given in Proposition 4.2 and Proposition 4.1.

Let us start with the bound given in Proposition 4.2. The partial balance condition (4.4) holds, and all birth and death processes $\tilde{Q}_i$, $i = 1,2,3$ are equal in distribution. Denote the arrival intensity of $\tilde{Q}_i$ process by $\tilde{\lambda}_i$ and its service rate by $\tilde{\mu}_i$. We have $\tilde{\lambda}_i = \lambda r_{0i} = \lambda/3$ and $\tilde{\mu}_i = \mu_i r_{i0} = c\lambda/3$. As already indicated in the introduction the equation for the L^2 spectral gap, for ergodic birth and death processes with constant rates, is known. The L^2 spectral gap (and the corresponding essential spectral gap) of $\tilde{Q}_i$ is given by

$$\operatorname{gap}(\tilde{Q}_i) = (\sqrt{\tilde{\mu}_i} - \sqrt{\tilde{\lambda}_i})^2 = \frac{\lambda}{3}(\sqrt{c}-1)^2,$$

therefore, the resulting bound is

$$\operatorname{gap}(\boldsymbol{Q}(\boldsymbol{Z})) \geq \frac{\lambda}{3}(\sqrt{c}-1)^2.$$

It is worth mentioning that this bound does not depend on p or q. Moreover,

$$\inf_{\substack{p,q\in(0,1)\\ p+q<1}} \operatorname{gap}_{\text{ess}} := \frac{\lambda}{3}(\sqrt{c}-1)^2.$$

On the other hand,

$$\sup_{\substack{p,q\in(0,1)\\ p+q<1}} \text{gap}_{\text{ess}} = \lambda(\sqrt{c}-1)^2,$$

which means that the bound given in Proposition 4.2 is at most three times smaller than the considered upper bound on the spectral gap. Moreover, $\text{gap}(\boldsymbol{Q}(\boldsymbol{Z}))$ is arbitrarily close to $(\lambda/3)(\sqrt{c}-1)^2$ for small values of p and q.

Now, let us turn to Proposition 4.1. Each distribution π_i is geometric with the corresponding hazard functions $h_{\pi_i}(n) = 1 - 1/c$. We have $r_{i0} > 0$ and $r_{0i} > 0$ for $i = 1, 2, 3$, thus, we can reduce $1 + \bar{d}\bar{b}(2m+1)$ to 1 in this proposition. We need yet to calculate:

$$|\boldsymbol{Q}| = \lambda r_{01} + \lambda r_{02} + \lambda r_{03} + \mu_1 + \mu_2 + \mu_3 = \lambda + 3c\frac{\lambda}{3(1-(p+q))} = \lambda\left(1 + \frac{c}{1-(p+q)}\right)$$

$$q^{\min} = \min\left(\frac{\lambda}{3}, \mu(1-(p+q)), \mu_i p, \mu_i q\right) = \frac{\lambda}{3}\min\left(1, \frac{cp}{1-(p+q)}, \frac{cq}{1-(p+q)}\right)$$

$$\check{q}^{\max} = 3\lambda_1 + 3\mu_1 = 3(1+c)\lambda_1 = \frac{\lambda(1+c)}{1-(p+q)}.$$

For the resulting bound with $\lambda = 1$, c ranging from 2 to 9 and for p and q close to 0, the ratio of the spectral gap and (4.3) in the best case is of order 10^{-5}. In this example the bound (4.3) is rather rough.

Example 6.2. Let $\boldsymbol{Z}$ be the classical *completely symmetrical* Jackson network with m stations, the routeing matrix R given by $r_{ij} = p < 1/(m-1)$ for all $i \neq j$, $r_{0i} = 1/m$, $i, j = 1, \dots, m$, and the arrival intensity λ. Note that we have $r_{i0} = 1 - (m-1)p$ for $i = 1, \dots, m$. The solution of the traffic equation is given by $\lambda_i = (1/m)(\lambda/1 - (m-1)p)$ for all $i = 1, \dots, m$. Moreover, assume that $\mu_i = c\lambda_i$, $c > 1$. Then the assumptions of Proposition 4.3(i) are fulfilled and

$$\begin{aligned}\text{gap}(\boldsymbol{Q}(\boldsymbol{Z})) &\leq \text{gap}_{\text{ess}} \\ &:= \left(1 - \frac{(m-1)p^2}{1-(m-2)p}\right)\lambda_i(\sqrt{c}-1)^2 \\ &= \frac{1}{m}\frac{1+p}{1-p(m-2)}(\sqrt{c}-1)^2\lambda.\end{aligned}$$

Note that for $p \in (0, 1/(m-1))$ we have

$$\frac{1}{m}(\sqrt{c}-1)^2\lambda \leq \text{gap}_{\text{ess}} \leq (\sqrt{c}-1)^2\lambda$$

Let us compare the value of the upper bound with the lower bound obtained in Proposition 4.2. Again, the partial balance condition (4.4) holds, and all birth and death processes $\tilde{Q}_i$, $i = 1, \dots, m$, are equal in distribution. The intensities are $\tilde{\lambda}_i = \lambda r_{0i} = \lambda/m$ and $\tilde{\mu}_i = \mu_i r_{i0} = c\lambda/m$. We have (similarly, as in the previous example)

$$\text{gap}(\tilde{Q}_i) = (\sqrt{\tilde{\mu}_i} - \sqrt{\tilde{\lambda}_i})^2 = \frac{\lambda}{m}(\sqrt{c}-1)^2,$$

therefore,

$$\text{gap}(\boldsymbol{Q}(\boldsymbol{Z})) \geq \frac{\lambda}{m}(\sqrt{c}-1)^2.$$

The obtained bound is the best we can have as a bound which is independent of p. The lower bound is at most m times smaller than the above given upper bound of the spectral gap. Moreover, the exact value $\mathrm{gap}(\boldsymbol{Q}(\boldsymbol{Z}))$ can be arbitrarily close to $(\lambda/m)(\sqrt{c}-1)^2$ for small values of p.

Regarding the bound from Proposition 4.1, again each π_i, $i=1,\ldots,m$, is geometric with the hazard function $h_{\pi_i}(n)=1-1/c$. We can reduce $1+\bar{d}\bar{b}(2m+1)$ to 1. We need to calculate the following constants:

$$|\boldsymbol{Q}| = \lambda\biggl(1+\frac{c}{1-(m-1)p}\biggr),$$

$$q^{\min} = \min\biggl(\frac{\lambda}{m}, \mu_i(1-(m-1)p), \mu_i p\biggr) = \frac{\lambda}{m}\min\biggl(1, \frac{cp}{1-(m-1)p}\biggr),$$

$$\check{q}^{\max} = m\lambda_1 + m\mu_1 = \frac{\lambda(1+c)}{1-(m-1)p}.$$

We skip writing the exact equation for the lower bound. The resulting values with $\lambda=1$, c ranging from 2 to 9 and for p close to 0, compared to the spectral gap, in the best case, are of order 10^{-5}, so the bound (4.3) is again rather rough.

Remark 6.1. Although the bounds obtained from our Proposition 4.1 gave rather rough results, it is worth stressing that it is possible to compute them for a large class of networks with variable service rates and unreliable nodes. The results that are possible to obtain via Proposition 4.3 are limited to very special cases of classical networks with constant service intensities. The bounds from Proposition 4.2 are limited to reliable networks and require a kind of partial balance (4.4) (which is fulfilled, for example, for reversible networks) but they are applicable to networks with variable service intensities and seem to work quite well. It is not true in general that the gap for a network is equal to the gap of a bottleneck station in this network. There still remains a lot of research to do in order to provide good computable bounds for networks, especially when the service rates are dependent on the queue size and where the nodes can be unreliable.

Acknowledgement

The authors are grateful to the referee for valuable comments which improved this paper.

References

[1] Anantharam, V. (1989). Threshold phenomena in the transient behaviour of Markovian models of communication networks and databases. *Queueing Systems Theory Appl.* **5,** 77–98.

[2] Baxendale, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Prob.* **15,** 700–738.

[3] Berenhaut, K. S. and Lund, R. (2001). Geometric renewal convergence rates from hazard rates. *J. Appl. Prob.* **38,** 180–194.

[4] Berenhaut, K. S. and Lund, R (2002). Renewal convergence rates for DHR and NWU lifetimes. *Prob. Eng. Inf. Sci.* **16,** 67–84.

[5] Blanc, J. P. C. (1985). The relaxation time of two queueing systems in series. *Commun. Statist. Stoch. Models* **1,** 1–16.

[6] Callaert, H. and Keilson, J. (1973). On exponential ergodicity and spectral structure for birth–death processes. I. *Stoch. Process. Appl.* **1,** 187–216.

[7] Chafai, D. and Joulin, A. (2013). Intertwining and commutation relations for birth–death processes. *Bernoulli* **19,** 1855–1879.

[8] Chen, M.-F. (1991). Exponential L^2-convergence and L^2-spectral gap for Markov processes. *Acta Math. Sinica (N.S.)* **7,** 19–37.

[9] CHEN, M.-F. (1996). Estimation of spectral gap for Markov chains. *Acta Math. Sinica (N.S.)* **12,** 337–360.

[10] CHEN, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.

[11] CHEN, M.-F. (2010). Speed of stability for birth–death processes. *Front. Math. China* **5,** 379–515.

[12] DADUNA, H. AND SZEKLI, R. (1995). Dependencies in Markovian networks. *Adv. Appl. Prob.* **27,** 226–254.

[13] DADUNA, H. AND SZEKLI, R. (1996). A queueing theoretical proof of increasing property of Pólya frequency functions. *Statist. Prob. Lett.* **26,** 233–242.

[14] DADUNA, H. AND SZEKLI, R. (2008). Impact of routeing on correlation strength in stationary queueing network processes. *J. Appl. Prob.* **45,** 846–878.

[15] DADUNA, H., KULIK, R., SAUER, C. AND SZEKLI, R. (2006). Dependence ordering for queueing networks with breakdown and repair. *Prob. Eng. Inf. Sci.* **20,** 575–594.

[16] DIACONIS, P. AND FILL, J. A. (1990). Examples for the theory of strong stationary duality with countable state spaces. *Prob. Eng. Inf. Sci.* **4,** 157–180.

[17] DIACONIS, P. AND FILL, J. A. (1990). Strong stationary times via a new form of duality. *Ann. Prob.* **18,** 1483–1522.

[18] DIACONIS, P. AND MICLO, L. (2009). On times to quasi-stationarity for birth and death processes. *J. Theoret. Prob.* **22,** 558–586.

[19] DIACONIS, P. AND STROOCK, D. (1991). Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Prob.* **1,** 36–61.

[20] DIEKER, A. B. AND WARREN, J. (2010). Series Jackson networks and noncrossing probabilities. *Math. Operat. Res.* **35,** 257–266.

[21] FAYOLLE, G., MALYSHEV, V. A., MEN´SHIKOV, M. V. AND SIDORENKO, A. F. (1993). Lyapounov functions for Jackson networks. *Math. Operat. Res.* **18,** 916–927.

[22] FILL, J. A. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Prob.* **1,** 62–87.

[23] FILL, J. A. (1991). Time to stationarity for a continuous-time Markov chain. *Prob. Eng. Inf. Sci.* **5,** 61–76.

[24] FILL, J. A. (1992). Strong stationary duality for continuous-time Markov chains. I. Theory. *J. Theoret. Prob.* **5,** 45–70.

[25] FILL, J. A. (2009). On hitting times and fastest strong stationary times for skip-free and more general chains. *J. Theoret. Prob.* **22,** 587–600.

[26] HART, A. G., MARTÍNEZ S. AND SAN MARTIN J. (2003). The λ-classification of continuous-time birth-and-death processes. *Adv. Appl. Prob.* **35,** 1111–1130.

[27] HORDIJK, A. AND SPIEKSMA, F. (1992). On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Prob.* **24,** 343–376.

[28] IGNATIOUK-ROBERT, I. (2006). On the spectrum of Markov semigroups via sample path large deviations. *Prob. Theory Relat. Fields* **134,** 44–80.

[29] IGNATIOUK-ROBERT, I. AND TIBI, D. (2012). Explicit Lyapunov functions and estimates of the essential spectral radius for Jackson networks. Preprint. Available at http://arxiv.org/abs/1206.3066v1.

[30] ISCOE, I. AND MCDONALD, D. (1994). Asymptotics of exit times for Markov jump processes. II. Applications to Jackson networks. *Ann. Prob.* **22,** 2168–2182.

[31] KARTASHOV, N. V. (2000). Determination of the spectral index of ergodicity of a birth-and-death process. *Ukrainian Math. J.* **52,** 1018–1028.

[32] KIJIMA M. (1992). Evaluation of the decay parameter for some specialized birth–death processes. *J. Appl. Prob.* **29,** 781–791.

[33] KROESE, D. P., SCHEINHARDT, W. R. W. AND TAYLOR, P. G. (2004). Spectral properties of the tandem Jackson network, seen as a quasi-birth-and-death process. *Ann. Appl. Prob.* **14,** 2057–2089.

[34] LAWLER, G. F. AND SOKAL, A. D. (1988). Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309,** 557–580.

[35] LIGGETT, T. M. (1989). Exponential L_2 convergence of attractive reversible nearest particle systems. *Ann. Prob.* **17,** 403–432.

[36] LIU, W. AND MA, Y. (2009). Spectral gap and convex concentration inequalities for birth–death processes. *Ann. Inst. H. Poincaré Prob. Statist.* **45,** 58–69.

[37] LOREK, P. AND SZEKLI, R. (2012). Strong stationary duality for Möbius monotone Markov chains. *Queueing Systems* **71,** 79–95.

[38] LUND, R. B., MEYN, S. P. AND TWEEDIE, R. L. (1996). Computable exponential convergence rates for stochastically ordered Markov processes. *Ann. Appl. Prob.* **6,** 218–237.

[39] MALYSHEV, V. A. AND SPIEKSMA, F. M. (1995). Intrinsic convergence rate of countable Markov chains. *Markov Processes Relat. Fields* **1,** 203–266.

[40] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability.* Springer, London.

[41] ROBERTS, G. O. AND TWEEDIE, R. L. (2000). Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Prob.* **37,** 359–373.

[42] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley, Chichester.
[43] Sauer, C. (2006). Stochastic product form networks with unreliable nodes: analysis of performance and availability. Doctoral Thesis, Hamburg University.
[44] Sauer, C. and Daduna, H. (2003). Availability formulas and performance measures for separable degradable networks. *Econom. Quality Control* **18,** 165–194.
[45] Sirl, D., Zhang, H. and Pollett, P. (2007). Computable bounds for the decay parameter of a birth–death process. *J. Appl. Prob.* **44,** 476–491.
[46] Van Doorn, E. A. (1981). *Stochastic Monotonicity and Queueing Applications of Birth–Death Processes* (Lecture Notes Statis. **4**). Springer, New York.
[47] Van Doorn, E. A. (1985). Conditions for exponential ergodicity and bounds for the decay parameter of a birth–death process. *Adv. Appl. Prob.* **17,** 514–530.
[48] Van Doorn, E. A. (2002). Representations for the rate of convergence of birth–death processes. *Theory Prob. Math. Statist.* **65,** 37–43.
[49] Van Doorn, E. A., Zeifman, A. I. and Panfilova, T. L. (2010). Bounds and asymptotics for the rate of convergence of birth–death processes. *Theory Prob. Appl.* **54,** 97–113.
[50] Vere-Jones, D. (1963). On the spectra of some linear operators associated with queueing systems. *Z. Wahrscheinlichkeitsth.* **2,** 12–21.
[51] Wu, L. (2004). Essential spectral radius for Markov semigroups. I. Discrete time case. *Prob. Theory Relat. Fields* **128,** 255–321.