Stochastic Bounds for Fast Jackson Networks

James Martin[1]

University of Cambridge

Abstract

We consider a Jackson-type network, each of whose nodes contains N identical channels with a single server. Upon arriving at a node, a task selects m of the channels at random, and enters the one with the shortest queue. We provide stochastic bounds for the distribution of the state of the network in terms of the behaviour of a collection of independent nodes of the same type. (Such nodes in isolation were introduced by Vvedenskaya, Dobrushin and Karpelevich). The bounds are asymptotically tight when every node is at least fully loaded. This generalises results of Massey, who gave bounds for the behaviour of a standard Jackson network in terms of that of a collection of independent $M/M/1$ queues.

1 Introduction

In their ground-breaking paper [7], Vvedenskaya, Dobrushin and Karpelevich consider a model of a service station consisting of several separate servers (see also [6] and [4]). The system is Markovian and consists of N identical channels, each with an infinite buffer and a single server with service rate μ. The input flow is a Poisson process with rate $N\lambda$. Upon arrival, each task selects m channels uniformly at random (with replacement), and joins the one with the shortest queue (breaking ties at random if necessary). We will call such a system a *VDK node.*

We will consider Markovian networks of Jackson type whose nodes behave like VDK nodes. Each node $j \in \{1, \ldots, J\}$ of the network contains N channels, each of which has an infinite buffer and a single exponential server with service rate μ_j. The network is fed by a family of Poisson flows of rates $N\lambda_1, \ldots, N\lambda_J$ arriving at the corresponding nodes. After being served at node j, a task jumps to node k with probability p_{jk}, and leaves the network with probability $p_j^* = 1 - \sum_k p_{jk}$, where $0 \le p_{jk} \le 1$ for all j, k and $\sum_k p_{jk} \le 1$ for all j. Upon arrival at node j, either from outside the network or from another node, a task selects m of the N channels there uniformly at random, and joins the one with the shortest queue. (Again, we will assume that the selection is made with replacement, though it could be made without replacement if $m \le N$). All arrival processes, service times and routing decisions are mutually independent. The network is therefore specified by the parameters J, N, $\underline{\lambda}$, $\underline{\mu}$, $\mathbf{P}$ and m. (Throughout the paper, underlined symbols will refer to J-component vectors).

[1]Correspondence address: Jesus College, Cambridge CB5 8BL, UK.
Email: `J.B.Martin@statslab.cam.ac.uk`

Such networks are studied in [2]. There it is shown that, as $N \to \infty$, the evolution of the system may be asymptotically represented by the solution of a countably infinite system of ODEs; under a standard non-overload condition on the parameters $\underline{\lambda}$, $\underline{\mu}$ and $\mathbf{P}$, the system has an invariant distribution π_N for each N, and, as $N \to \infty$, π_N converges to a limiting invariant distribution which is concentrated at a single point, corresponding to the fixed point of the system of ODEs. Under this limiting distribution, the tail of the distribution of queue lengths decays super-exponentially, rather than exponentially as in the case of a standard Jackson network — hence the term "fast Jackson network".

In this paper we provide stochastic bounds on the behaviour of the network which apply also in the case where the non-overload condition does not hold, and the system is transient. This generalises a result of Massey [3] (see also [5] and the remarks after the proof of Theorem 1 below), who gave bounds on the evolution of a standard Jackson network in terms of the behaviour of a system of independent $M/M/1$ queues, whose arrival and service rates are chosen according to those of the network. Our Theorem 1 gives analogous bounds for the evolution of a fast Jackson network in terms of the behaviour of a system of independent VDK nodes. Theorem 2 shows that the bounds are in some sense "best possible" when every node of the network is at least fully loaded.

The bounds are expressed in terms of a partial ordering between distributions of network states. This ordering compares the probabilities of certain increasing sets. In general there is no corresponding pathwise, or even pointwise, coupling which gives an almost sure ordering of the states of the two systems (see the remarks after Theorem 1).

2 Stochastic Bounds

We describe the state of a network with J N-channel nodes at time t by a vector $\underline{\mathbf{X}}(t)$ $= \{X_{j,r}(t),\, 1 \le j \le J,\, 1 \le r \le N\} \in (\mathbb{Z}_+^N)^J$ (where $\mathbb{Z}_+$ is the set of non-negative integers); $X_{j,r}(t)$ represents the length of the queue in the rth channel of the jth node at time t.

Call $\Gamma \subset \mathbb{Z}_+^N$ an *increasing set* if, whenever $\mathbf{x} \in \Gamma$ and $\mathbf{x'} \ge \mathbf{x}$, we have $\mathbf{x'} \in \Gamma$ also. For random variables $\underline{\mathbf{X}}, \underline{\mathbf{Y}}$ taking values in $(\mathbb{Z}_+^N)^J$, write $\underline{\mathbf{X}} \ge_d \underline{\mathbf{Y}}$ if

$$\mathbb{P}(\mathbf{X}_1 \in \Gamma_1, \dots, \mathbf{X}_J \in \Gamma_J) \ge \mathbb{P}(\mathbf{Y}_1 \in \Gamma_1, \dots, \mathbf{Y}_J \in \Gamma_J)$$

whenever $\Gamma_1, \dots, \Gamma_J$ are increasing sets in $\mathbb{Z}_+^N$.

Note that in the case of a standard Jackson network, one may freely assume that $p_{ii} = 0$ for all i, since a departure which immediately reenters the same node does not change the state of the system. Here, this is not the case, since the task may leave one channel and join another at the same node – so we extend slightly the model of a VDK node considered in [7] to include the possibility that some departures from the system are immediately reintroduced (and join the shortest of m randomly selected queues in the normal way, independently of previous choices).

The following theorem gives the main upper bound on the behaviour of the network, and also a corresponding lower bound in terms of the behaviour of a single $M/M/1$ queue.

Theorem 1 *Let $\underline{\mathbf{Y}}(t)$ be a fast Jackson network with parameters $J, N, m, \underline{\lambda}, \underline{\mu}, \mathbf{P}$.*

(a) Let $\underline{\mathbf{X}}(t)$ be a system of J independently evolving VDK nodes, with parameters N and m, arrivals at node i at rate $N(\lambda_i + \sum_{j \neq i} \mu_j p_{ji})$ and service at node i at rate μ_i, and with a proportion p_{ii} of departures at node i immediately reintroduced to the same node. Suppose that the time 0 distributions of $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ satisfy $\underline{\mathbf{X}}(0) \geq_d \underline{\mathbf{Y}}(0)$. Then $\underline{\mathbf{X}}(t) \geq_d \underline{\mathbf{Y}}(t)$ for all $t \geq 0$.

(b) Let $Z(t)$ be a single $M/M/1$ queue, with arrival rate $N \sum_i \lambda_i$ and service rate $N \sum_i \mu_i p_i^*$, and with

$$\mathbb{P}(Z(0) \geq z) \leq \mathbb{P}\left(\sum_{i=1}^{J} \sum_{r=1}^{N} Y_{i,r}(0) \geq z \right)$$

for all z. Then, for all t, z,

$$\mathbb{P}(Z(t) \geq z) \leq \mathbb{P}\left(\sum_{i=1}^{J} \sum_{r=1}^{N} Y_{i,r}(t) \geq z \right).$$

Proof:

(a) We will couple the evolution of $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$. Consider a collection of Poisson processes A and $\tilde{A}$, all independent of each other, indexed variously by $i, j \in \{0, \ldots, J\}$ and $a, b_1, b_2, \ldots, b_m \in \{1, \ldots, N\}$. The indices have the following meanings, where relevant: i and a represent the node and channel respectively from which a task departs; j represents the node to which a task arrives; and $b_1, \ldots, b_m$ represent the channels selected on arrival at the node — the task joins the one with the shortest queue (choosing the first one listed in the case of a tie). If $i = 0$ (respectively $j = 0$), this indicates an arrival from (respectively a departure to) outside the network. The word "potential" below means that the transfer or departure referred to only occurs if the channel from which a task is to depart is non-empty. The collection of processes may be divided into five types:

(i) $A_{i,a,j,b_1,\ldots,b_m}, i, j \geq 1, i \neq j$, of rates $\mu_i p_{ij} N^{-m}$, whose points represent, simultaneously, potential transfers from node i to node j in $\underline{\mathbf{Y}}$ and potential departures from node i in $\underline{\mathbf{X}}$.

(ii) $\tilde{A}_{i,a,j,b_1,\ldots,b_m}, i, j \geq 1, i \neq j$, of rates $\mu_i p_{ij} N^{-m}$, whose points represent arrivals at node j in $\underline{\mathbf{X}}$. (Here the nature of the event is not affected by i and a, but the indices are retained for comparison with type (i)).

(iii) $A_{i,a,i,b_1,\ldots,b_m}, i \geq 1$, of rates $\mu_i p_{ii} N^{-m}$ representing, in both $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$, potential transfers within node i.

(iv) $A_{0,0,j,b_1,\ldots,b_m}, j \geq 1$, of rates $\lambda_j N^{-m+1}$ representing, in both $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$, arrivals at node j.

(v) $A_{i,a,0,0,\ldots,0}, i \geq 1$, of rates $\mu_i p_i^*$ representing, in both $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$, potential departures from node i.

Under the coupling thereby described, $\underline{\mathbf{X}}(t)$ and $\underline{\mathbf{Y}}(t)$ are Markov processes of the required types, and the pair $(\underline{\mathbf{X}}(t), \underline{\mathbf{Y}}(t), t \geq 0)$ forms a Markov process, whose jump times are a subset of the points of the various processes A and $\tilde{A}$.

Let $T_0 = 0$ and, for $n \geq 1$, let T_n be the time of the nth point after time 0 of all the processes $A, \tilde{A}$. Write $\underline{\mathbf{X}}^{(n)} = \underline{\mathbf{X}}(T_n+), \underline{\mathbf{Y}}^{(n)} = \underline{\mathbf{Y}}(T_n+)$. For appropriate i, a, j, $b_1, \ldots, b_m$, say that the event $G_n(i, a, j, b_1, \ldots, b_m)$ occurs if the point at T_n is a point in a process $A_{i,a,j,b_1,\ldots,b_m}$ or $\tilde{A}_{i,a,j,b_1,\ldots,b_m}$.

Since all the $A, \tilde{A}$ are independent and Poisson, the points $\{T_n, n \geq 1\}$ form a Poisson process; the process $\{T_n, n \geq 1\}$ is independent of the events $\{G_n, n \geq 1\}$ and so also of $\{\underline{\mathbf{X}}^{(n)}, \underline{\mathbf{Y}}^{(n)}, n \geq 0\}$. Also for any n', the events $\{G_n, n > n'\}$ are independent of the collection $\{\underline{\mathbf{X}}^{(n)}, \underline{\mathbf{Y}}^{(n)}, n \leq n'\}$. Now

$$\begin{aligned}
\mathbb{P}(\underline{\mathbf{X}}(t) \in \Gamma_1 \times \ldots \times \Gamma_J) &= \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq t < T_{n+1}, \underline{\mathbf{X}}^{(n)} \in \Gamma_1 \times \ldots \times \Gamma_J) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq t < T_{n+1}) \mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1 \times \ldots \times \Gamma_J),
\end{aligned}$$

and similarly for $\underline{\mathbf{Y}}(t)$, so it will be enough to show that $\underline{\mathbf{X}}^{(n)} \geq_d \underline{\mathbf{Y}}^{(n)}$ for all $n > 0$. By assumption, $\underline{\mathbf{X}}^{(0)} \geq_d \underline{\mathbf{Y}}^{(0)}$, so by induction it suffices to show that if $\underline{\mathbf{X}}^{(n)} \geq_d \underline{\mathbf{Y}}^{(n)}$, then also $\underline{\mathbf{X}}^{(n+1)} \geq_d \underline{\mathbf{Y}}^{(n+1)}$.

We have

$$\begin{aligned}
\mathbb{P}(\underline{\mathbf{X}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J) = \sum_{i,j,a,b_1,\ldots,b_m} &\mathbb{P}(G_{n+1}(i, a, j, b_1, \ldots, b_m)) \\
&\times \mathbb{P}(\underline{\mathbf{X}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J \,\big|\, G_{n+1}(i, a, j, b_1, \ldots, b_m)),
\end{aligned}$$

and similarly for $\underline{\mathbf{Y}}^{(n+1)}$. So in fact it will be enough to show that if $\underline{\mathbf{X}}^{(n)} \geq_d \underline{\mathbf{Y}}^{(n)}$ then, for all $\Gamma_1, \ldots, \Gamma_J$ increasing, all appropriate $i, j, a, b_1, \ldots, b_m$,

$$\begin{aligned}
\mathbb{P}(\underline{\mathbf{X}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J &\,\big|\, G_{n+1}(i, a, j, b_1, \ldots, b_m)) \\
&\geq \mathbb{P}(\underline{\mathbf{Y}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J \,\big|\, G_{n+1}(i, a, j, b_1, \ldots, b_m)). \quad (2.1)
\end{aligned}$$

We make the following observation. Fix $a, b_1, \ldots, b_m$, and take an increasing set $\Gamma_k \subset \mathbb{Z}_+^N$ describing a collection of states at node k. Define the set Γ_k^+ of states at node k which, after a departure from channel a, yield a state in Γ_k. Then Γ_k^+ is also an increasing set, and $\Gamma_k^+ \subset \Gamma_k$. Similarly, define the set Γ_k^- of states which, after an arrival at the channel out of $b_1, \ldots, b_m$ which has the shortest queue, yield a state in Γ_k; then Γ_k^- is increasing and $\Gamma_k \subset \Gamma_k^-$. Finally, consider a transfer from channel a back to the same node, joining the shortest queue from $b_1, \ldots, b_m$; the set $\Gamma_k^=$ of states from which this transfer yields a state in Γ_k is increasing.

We return to (2.1). First suppose that $i = j$, so that we have a point in a process of type (iii). Without loss of generality, take $i = 1$. Then

$$\begin{aligned}
\mathbb{P}(\underline{\mathbf{X}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J &\,\big|\, G_{n+1}(1, a, 1, b_1, \ldots, b_m)) \\
&= \mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1^= \times \Gamma_2 \times \ldots \times \Gamma_J \,\big|\, G_{n+1}(1, a, 1, b_1, \ldots, b_m)) \\
&= \mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1^= \times \Gamma_2 \times \ldots \times \Gamma_J) \\
&\geq \mathbb{P}(\underline{\mathbf{Y}}^{(n)} \in \Gamma_1^= \times \Gamma_2 \times \ldots \times \Gamma_J) \\
&= \mathbb{P}(\underline{\mathbf{Y}}^{(n)} \in \Gamma_1^= \times \Gamma_2 \times \ldots \times \Gamma_J \,\big|\, G_{n+1}(1, a, 1, b_1, \ldots, b_m)) \\
&= \mathbb{P}(\underline{\mathbf{Y}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J \,\big|\, G_{n+1}(1, a, 1, b_1, \ldots, b_m))
\end{aligned}$$

as required. The cases $i = 0$ (type (iv)) and $j = 0$ (type (v)) are similar.

If $i, j \geq 1$ and $i \neq j$, then given $G_{n+1}(i, a, j, b_1, \dots, b_m)$ we have one out of type (i) and type (ii), with equal probability. Without loss of generality, let $i = 1$ and $j = 2$. Then

$$
\begin{aligned}
&\mathbb{P}(\underline{\mathbf{X}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J \,|\, G_{n+1}(1, a, 2, b_1, \dots, b_m)) \\
&\quad = \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1^+ \times \Gamma_2 \times \Gamma_3 \times \ldots \times \Gamma_J) + \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1 \times \Gamma_2^- \times \Gamma_3 \times \ldots \times \Gamma_J) \qquad (2.2)\\
&\quad \geq \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1^+ \times \Gamma_2^- \times \Gamma_3 \times \ldots \times \Gamma_J) + \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{X}}^{(n)} \in \Gamma_1 \times \Gamma_2 \times \Gamma_3 \times \ldots \times \Gamma_J) \qquad (2.3)\\
&\quad \geq \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{Y}}^{(n)} \in \Gamma_1^+ \times \Gamma_2^- \times \Gamma_3 \times \ldots \times \Gamma_J) + \tfrac{1}{2}\mathbb{P}(\underline{\mathbf{Y}}^{(n)} \in \Gamma_1 \times \Gamma_2 \times \Gamma_3 \times \ldots \times \Gamma_J)\\
&\quad = \mathbb{P}(\underline{\mathbf{Y}}^{(n+1)} \in \Gamma_1 \times \ldots \times \Gamma_J \,|\, G_{n+1}(1, a, 2, b_1, \dots, b_m)).
\end{aligned}
$$

Here, to get from (2.2) to (2.3), we use the fact that if $C \subseteq C'$ and $D \subseteq D'$, then

$$\mathbb{P}(C \cap D') + \mathbb{P}(C' \cap D) \leq \mathbb{P}(C \cap D) + \mathbb{P}(C' \cap D').$$

Thus (2.1) holds for all i, j, which completes the proof of part (a).

(b) One can construct a simple coupling under which $Z(t) \leq \sum_{i=1}^{J} \sum_{r=1}^{N} Y_{i,r}(t)$ for all t. Let arrivals to Z occur at the same time as external arrivals to any node in $\underline{\mathbf{Y}}$, and let potential departures from Z occur at the same time as potential departures to outside the network from any of the channels in the nodes of $\underline{\mathbf{Y}}$. Departures from Z are only lost if the queue is already empty, while departures in $\underline{\mathbf{Y}}$ are lost whenever the selected channel is empty, even if other channels at the same or another node are not. $\square$

Remarks:

A stochastic domination of the sort given in Theorem 1(a) by *non-independent* VDK nodes follows from a simple coupling (with a system where all the nodes output tasks even when empty). The more intricate argument above is needed to provide a dominating system whose nodes are independent.

The partial order $\geq_d$ on the set of probability distributions on $(\mathbb{Z}_+^N)^J$ is not easily interpretable in terms of the usual partial order on the set $(\mathbb{Z}_+^N)^J$ itself. This is illustrated by the construction of Massey in [3]. Firstly, one cannot in general make a pathwise coupling of the processes $\underline{\mathbf{X}}(t)$ and $\underline{\mathbf{Y}}(t)$, $t \geq 0$ in such a way that $\mathbb{P}(\underline{\mathbf{X}}(t) \geq \underline{\mathbf{Y}}(t) \text{ for all } t \geq 0) = 1$ (this is also observed in Remark (3) in [5]). But more strongly, even for fixed t one cannot generally couple the random variables $\underline{\mathbf{X}}(t)$ and $\underline{\mathbf{Y}}(t)$ in such a way that $\mathbb{P}(\underline{\mathbf{X}}(t) \geq \underline{\mathbf{Y}}(t)) = 1$ (for the case $N = 1$, this is precisely Theorem 3.2 of [3]). Similarly the random variables $\underline{\mathbf{X}}^{(n)}$ and $\underline{\mathbf{Y}}^{(n)}$ (defined in the proof of Theorem 1 above) cannot generally be coupled in such a way that $\mathbb{P}\big(\underline{\mathbf{X}}^{(n)} \geq \underline{\mathbf{Y}}^{(n)}\big) = 1$ — the same method of proof as for Theorem 3.2 of [3] applies.

Putting $N = 1$ reduces to the case of a standard Jackson network and $M/M/1$ queues. Then any increasing set Γ of the type considered has the simple form $\{x : x \geq x_0\}$ for some $x_0 \in \mathbb{Z}_+$. The result of (1) in this case was proved by Massey in [3] using operator methods. A simpler probabilistic proof was proposed by Tsoucas and Walrand in [5], and our proof of Theorem 1(a) above draws extensively on their coupling method. However, it is not clear that the proof in [5] is complete; the induction step used there seems to rely on the sort of pointwise coupling between $\underline{X}^{(n)}$ and $\underline{Y}^{(n)}$ described above, which is not generally available.

An alternative reading of the proof would require the following: if $\underline{X}^{(n)} \geq_d \underline{Y}^{(n)}$ and $(X_1^{(n+1)}, X_2^{(n+1)}) \geq_d (Y_1^{(n+1)}, Y_2^{(n+1)})$, with $(X_i^{(n)}, Y_i^{(n)}) = (X_i^{(n+1)}, Y_i^{(n+1)})$ for $3 \leq i \leq J$, and with $(X_1^{(n+1)}, X_2^{(n+1)}, Y_1^{(n+1)}, Y_2^{(n+1)})$ independent of $(X_3^{(n)}, \ldots, X_J^{(n)}, Y_3^{(n)}, \ldots, Y_J^{(n)})$ given $(X_1^{(n)}, X_2^{(n)}, Y_1^{(n)}, Y_2^{(n)})$, then $\underline{X}^{(n+1)} \geq_d \underline{Y}^{(n+1)}$. But this is not true in general.

The proof given here (which may of course be written more straightforwardly in the case $N=1$) avoids these problems by considering more explicitly the influence of $(\underline{\mathbf{X}}^{(n)}, \underline{\mathbf{Y}}^{(n)})$ on $(\underline{\mathbf{X}}^{(n+1)}, \underline{\mathbf{Y}}^{(n+1)})$, by means of the sets Γ^+, Γ^-, $\Gamma^=$.

3 Tightness of Bounds

In this section, we give a generalisation of Theorem 3.6 of [3].

Define $\tilde{\rho}_i = \lambda_i + \sum_{j=1}^J \mu_j p_{ji}$. We will show that the bounds in Theorem 1 are asymptotically tight (in a sense to be made precise in Theorem 2) if $\tilde{\rho}_i \geq \mu_i$ for all i; equivalently, if $\rho_i \geq \mu_i$ for all i, where $\{\rho_i, 1 \leq i \leq J\}$ is the unique solution to the "throughput equations" $\rho_i = \lambda_i + \sum_{j=1}^J \min(\mu_j, \rho_j) p_{ji}$ (see [1] for details); thus every node in the network is working at full capacity. If $\tilde{\rho}_j < \mu_j$ for some j, then (with probability 1) the proportion of time for which a typical channel at node j is busy does not converge to 1 in the long run, so that the terms $\mu_j p_{ji}$, representing the long-run average rate of transfer from node j to node i, are overestimates, and the bound is not optimal (unless $p_j^* = 1$).

Theorem 3 of [2] compares the behaviour of a fast Jackson network to that of a "linear" network where $m = 1$; the following lemma is a special case where $J = 1$. (Alternatively, Theorem 4 of [6] provides an easier route to more sophisticated comparisons between VDK nodes with different values of m, though a small change to the coupling method used there is needed if some departures are to be reintroduced to the node).

Lemma 1 *Let $\mathbf{V}(t)$ and $\tilde{\mathbf{V}}(t)$ be VDK nodes with the same number of channels N, the same arrival rate, the same service rate, and the same proportion of departures immediately reintroduced to the node. Let arrivals in $\mathbf{V}$ inspect m channels, and arrivals in $\tilde{\mathbf{V}}$ inspect only one channel. If $\mathbf{V}(0)$ and $\tilde{\mathbf{V}}(0)$ have the same distribution, then there exists a coupling between the two systems such that, for all t, $\sum_{r=1}^N V_r(t) \leq \sum_{r=1}^N \tilde{V}_r(t)$.*

We can now prove the theorem on tightness of bounds, showing that the asymptotic growth rates of the queues in the fast Jackson network $\underline{\mathbf{Y}}$ coincide with those in the bounding systems $\underline{\mathbf{X}}$ and Z.

Theorem 2 *Let $\underline{\mathbf{X}}$, $\underline{\mathbf{Y}}$ and Z be as in Theorem 1. Assume that $\mathbb{E}[X_{i,r}(0)] < \infty$ for all i, r. Then if $\tilde{\rho}_i \geq \mu_i$ for all i,*

$$\lim_{t\to\infty} t^{-1}\mathbb{E}[Z(t)] = N \sum_i (\tilde{\rho}_i - \mu_i),$$

and, for all i and r,

$$\lim_{t\to\infty} t^{-1}\mathbb{E}[Y_{i,r}(t)] = \lim_{t\to\infty} t^{-1}\mathbb{E}[X_{i,r}(t)] = \tilde{\rho}_i - \mu_i.$$

Proof: For an $M/M/1$ queue $Q(t)$ with arrival rate α, service rate $\beta \leq \alpha$ and $\mathbb{E}[Q(0)] < \infty$,

$$\lim_{t\to\infty} t^{-1}\mathbb{E}[Q(t)] = \alpha - \beta. \tag{3.1}$$

Hence

$$\lim_{t\to\infty} t^{-1}\mathbb{E}[Z(t)] = N\sum_i \lambda_i - N\sum_i \mu_i p_i^*,$$
$$= N\sum_i (\tilde{\rho}_i - \mu_i), \tag{3.2}$$

as required. From Theorem 1, we also have, for all i, r, t,

$$\mathbb{E}[Y_{i,r}(t)] \le \mathbb{E}[X_{i,r}(t)] \tag{3.3}$$

and

$$\mathbb{E}[Z(t)] \le \sum_{i=1}^{J}\sum_{r=1}^{N} \mathbb{E}[Y_{i,r}(t)]. \tag{3.4}$$

Consider a system $\underline{\mathbf{X}}'(t)$ evolving in the same way as $\underline{\mathbf{X}}(t)$, but with initial distribution given by

$$\mathbb{P}(X'_{i,1}(0) = X'_{i,2}(0) = \ldots = X'_{i,N}(0)) = 1$$

for all i, and, for all i, r, z,

$$\mathbb{P}(X'_{i,r}(0) = z) = \mathbb{P}(\max_{1\le r'\le N} X_{i,r'}(0) = z).$$

Then one may easily construct a coupling between $\underline{\mathbf{X}}$ and $\underline{\mathbf{X}}'$ under which $\underline{\mathbf{X}}(t) \le \underline{\mathbf{X}}'(t)$ for all t, so that, for all i, r, t,

$$\mathbb{E}[X_{i,r}(t)] \le \mathbb{E}[X'_{i,r}(t)]. \tag{3.5}$$

Next, let $\underline{\mathbf{X}}''$ be a system identical to $\underline{\mathbf{X}}'$, except that $m = 1$ in the system $\underline{\mathbf{X}}''$. By Lemma 1, there is a coupling between the two systems $\underline{\mathbf{X}}'$ and $\underline{\mathbf{X}}''$ such that, for all i and t, $\sum_{r=1}^{N} X'_{i,r}(t) \le \sum_{r=1}^{N} X''_{i,r}(t)$. Using the symmetry among the channels at a node of $\underline{\mathbf{X}}'$ or of $\underline{\mathbf{X}}''$, we then have, for all i, r, t,

$$\mathbb{E}[X'_{i,r}(t)] \le \mathbb{E}[X''_{i,r}(t)]. \tag{3.6}$$

Now note that $\underline{\mathbf{X}}''$ may be regarded as a collection of J independent standard Jackson networks, each with N nodes. In each of these networks, every node has external arrival rate $\lambda_i + \sum_{j\neq i} \mu_j p_{ji}$ and service rate μ_i; a departing task leaves the network with probability $1 - p_{ii}$, and transfers to any of the N nodes (including the one from which it has just departed) with probability p_{ii}/N.

Then we can apply Theorem 1(a) to give, for each i, r, x, t,

$$\mathbb{P}(X''_{i,r}(t) \ge x) \le \mathbb{P}(W_i(t) \ge x),$$

where W_i is an $M/M/1$ queue with arrival rate $\lambda_i + \sum_{j=1}^{J} \mu_j p_{ji} - \mu_i p_{ii}/N$, with service rate μ_i and with reintroduction probability p_{ii}/N, and with $W_i(0)$ distributed as $\max_{r'} X_{i,r'}(0)$. Equivalently, we can let W_i have service rate $\mu_i(1 - p_{ii}/N)$ and reintroduction probability 0.

We have $\mathbb{E}[W_i(0)] \leq \sum_{r'} \mathbb{E}[X_{i,r'}(0)] < \infty$, so, by (3.1),

$$\begin{aligned}\lim_{t\to\infty} t^{-1}\mathbb{E}[X''_{i,r}(t)] \leq \lim_{t\to\infty} t^{-1}\mathbb{E}[W_i(t)] \\ &= \lambda_i + \sum_{j=1}^{J} \mu_j p_{ji} - \mu_i p_i i/N - \mu_i(1 - p_{ii}/N) \\ &= \tilde{\rho}_i - \mu_i. \end{aligned} \tag{3.7}$$

Combining (3.2)-(3.7) gives the required result. □

References

[1] Goodman, J. B. and Massey, W. A., (1984) The Non-Ergodic Jackson Network. *J. Appl. Prob.* **21**, 860–869.

[2] Martin, J. B. and Suhov, Y. M., (1999) Fast Jackson Networks. *Ann. Appl. Prob.* **9**, 854–870.

[3] Massey, W. A., (1984) An Operator-Analytic Approach to the Jackson Network. *J. Appl. Prob.* **21**, 379–393.

[4] Mitzenmacher, M., (1996) The Power of Two Choices in Randomized Load Balancing. Ph.D. thesis, University of California, Berkeley.

[5] Tsoucas, P. and Walrand, J., (1984) A Note on Stochastic Bounds for Queueing Networks. *Adv. Appl. Prob.* **16**, 926–928.

[6] Turner, S. R. E., (1998) The Effect of Increasing Routing Choice on Resource Pooling. *Probability in the Engineering and Informational Sciences* **12**, 109–124.

[7] Vvedenskaya, N. D., Dobrushin, R. L. and Karpelevich, F. I., (1996) Queueing System with Selection of the Shortest of Two Queues: An Asymptotic Approach. *Problems of Information Transmission* **32**, 15–27.