Applied Probability Trust (9 March 2007)

APPROXIMATING QUEUES IN SLOWLY VARYING STATIONARY ENVIRONMENTS

ANTONIS DIMAKIS,* *University of California at Berkeley*

JEAN WALRAND,** *University of California at Berkeley*

Abstract

We provide linear approximations to the marginal distributions for a class of infinite-state continuous-time stationary Markov chains in slowly varying environents. We take an approach motivated by light-traffic approximations to stationary point processes, which permits us to consider general stationary environments. Under mild assumptions we show that Jackson networks with routing not affected by the environment, belong to this framework.

Keywords: perturbation analysis, singular perturbations, Jackson network, random environment

2000 Mathematics Subject Classification: Primary 60K37

Secondary 60K25

1. Introduction

We consider approximations to the stationary distribution of a Markov chain, whose transition matrix is controlled by a piecewise constant process, henceforth called the environment, only assumed to be stationary. The problem is motivated by the presence of variations in channel quality and traffic demand in communication networks. In general, the computation of this distribution is not tractable and one resorts to simulation. Even in the case of a Markovian environment, where simulation is not necessary, the solution of a large number of equations can be undesirable. (Note that certain simple cases can be solved very efficiently, as in [2].) When environment transitions are rare, Taylor-like expansions around the time-scale separation case are possible under Markovian assumptions. Singular perturbation analysis (see [8, 11] and

* Postal address: Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720, USA

references therein) is one such approach, where tools there are algebraic. Also, analytic methods have been applied as in [4]. These do not apply in the non-Markovian case.

Here, we consider a framework inspired by light-traffic approximations ([5, 6]) and weak convergence instead, in computing the constant and first derivative terms. Key to applying such tools is tightness and uniform convergence properties, which are established by exploiting monotonicity inherent in specific models, and queueing networks in particular. We use monotonicity of Jackson networks, and show that it belongs to this framework.

As is commonly the case in light-traffic approximations, the first derivative is determined by what happens around a single jump. Hence, the approximations we obtain do not depend on more detailed statistics of the environment, other that its transitions. Thus, the approximations are the same as the ones given by a Markov model. It is conceivable that higher-order terms in such an expansion can be obtained using factorial moment expansions [6]. Also note that as we do not rely on uniformization, transition rates are not required to be bounded as assumed in singular perturbation methods (e.g., see [8]).

The main results of this paper are Theorems 1 and 2. Theorem 1 in section 2 provides an approximation framework for general Markov chains. In section 3 we show that Jackson networks belong to this framework (Theorem 2). As a side result, we establish existence of a stationary law for varying Jackson networks. Even though the stability conditions are not necessary (see [7] for single queues), stronger conditions are not needed in obtaining the approximations.

2. Approximation result

Consider a family of continuous-time Markov chains with countable state space $\mathcal{Y}$, given by the transition matrices $\{Q^x, x \in \mathcal{X}\}$ where $\mathcal{X}$ is a countable set. Assume each Q^x is positive recurrent and let π^x be the corresponding invariant distribution.

Assume a family of probability spaces $((\Omega, \mathcal{F}, P_\epsilon), \epsilon > 0)$ exists, endowed with a measure preserving flow $\theta = (\theta^t, t \in \mathbb{R})$. (For more details on flow compatibility and the stationary framework for point processes, the reader is referred to [9].)

On $(\Omega, \mathcal{F})$ the environment process $X = (X_t, t \in \mathbb{R})$ is given, compatible with flow θ, taking values in $\mathcal{X}$ with right-continuous piecewise-constant paths P_ϵ-a.s. for all $\epsilon > 0$. Assume the jump times $-\infty < \ldots < T_{-1} < T_0 \leq 0 < T_1 < \ldots < +\infty$ of X_t, form a simple point process. Under P_1, these jumps have finite intensity λ, and under P_ϵ, the jump times are dilated by ϵ^{-1}, i.e.,

$$P_\epsilon((T_{-n}, \ldots, T_n) \in A_{-n} \times \ldots \times A_n) = P_1((\epsilon^{-1}T_{-n}, \ldots, \epsilon^{-1}T_n) \in A_{-n} \times \ldots \times A_n)\,,$$

for each $\epsilon > 0, n = 0, 1, \ldots$, and Borel sets $(A_m : m = -n, \ldots, n)$. The sequence of values right after the jumps is not affected by ϵ, i.e. $P_\epsilon(X_{T_m} = x_m, m = -n, \ldots, n)$ does not depend on ϵ. Let P_ϵ^0 be the Palm distribution (w.r.t. the jumps of X_t) corresponding to P_ϵ. Also, with $\mathcal{F}^X$ we denote the σ-field generated by X.

On the same space, a process $Y = (Y_t, t \in \mathbb{R})$ is given compatible with θ, for which under P_ϵ and conditionally on X it is a Markov chain with generator Q^{X_t}. That is, for any $f : \mathcal{Y} \to \mathbb{R}$ with $E_\epsilon|f(Y_u)| < \infty$ for all u,

$$t \mapsto f(Y_t) - f(Y_s) - \int_s^t Q^{X_u} f(Y_{u-})du$$

is a martingale for all $-\infty < s \leq t < \infty$, w.r.t. the filtration generated by $\mathcal{F}^X$ and $(Y_u : u \leq t)$ for all t, and w.r.t. P_ϵ.

We will use multiple versions of Q^{X_t} Markov chains started from different initial states. These couple when they hit the same state and evolve according to $g : \mathcal{X} \times \mathcal{Y} \times \mathbb{R}_+ \times \Omega \to \mathcal{Y}$, defined such that

(a) $g(x, y, 0, \omega) = y\,, \quad g(x, g(x, y, t, \omega), s, \theta^t\omega) = g(x, y, t + s, \omega)\,, \quad \forall s, t \geq 0,$

(b) $(t, \omega) \mapsto g(x, y, t, \omega)$ is a measurable stochastic process for all x, y.

(c) $\mathcal{F}^X$ is independent of $\mathcal{G}$, the σ-field generated by $(g(x, y, t, \omega), x \in \mathcal{X}, y \in \mathcal{Y}, t \geq 0)$. (Intuitively, ω is "noise" independent of X.)

(d) For all x, y, the process $(t, \omega) \mapsto g(x, y, t, \omega)$ is a Markov chain with generator Q^x.

An example of such a function g is given in section 3.

We will compare Y with another process Z which corresponds to "$\epsilon = 0$". Before we define Z, let us informally explain the idea behind this since it plays an important role. At the jumps of X, e.g. at T_0, Z is independently started according to the invariant distribution that corresponds to the new state of the environment, i.e. $\pi^{X_{T_0}}$. Between times T_0 and T_1, Z evolves according to the transition matrix $Q^{X_{T_0}}$. Hence, Z behaves as Y with the difference that equilibrium is reached "instantaneously" by the former. Notice that the marginal distribution of Z at stationarity, is the one which Y is expected to converge to, as $\epsilon \downarrow 0$. If Z and Y are defined in an appropriate way then they may couple before T_1. Thus, $E_\delta(f(Y_0) - f(Z_0))$ measures the "finite difference" of the marginal distributions of Y between cases $\epsilon = \delta$ and $\epsilon = 0$.

We now define Z. Let $V_t = V_0 \circ \theta^t \in \mathcal{Y}$ be a right-continuous piecewise-constant process whose jumps are those of X, where the r.v's $V_{T_i}, i \in \mathbb{Z}$ are conditionally independent and each distributed according to $\pi^{X_{T_i}}$, given $\mathcal{F}^X$. Let $\mathcal{F}^V$ be the σ-field generated by $V = (V_t, t \in \mathbb{R})$. Define,

$$Z_t(\omega) = g(X_{T_0}, V_{T_0}, -T_0, \omega) \circ \theta^t \ , \quad t \in \mathbb{R} \, .$$

Now fix any $f : \mathcal{Y} \to \mathbb{R}$ and consider the assumptions:

(A1) $(P^0_\epsilon(Y_0 \in \cdot), \epsilon > 0)$ is tight.

(A2) $\inf\{t > 0 | g(X_0, Y_0, t, \omega) = g(X_0, V_0, t, \omega)\} < \infty \ , \quad P^0_\epsilon$-a.s., $\forall \epsilon > 0$.

(A3) $E_\epsilon |f(Y_0)| < \infty$, $E_\epsilon |f(V_0)| < \infty$.

(A4) There exist r.v. B for which

$$\int_0^\infty |f(g(X_0, Y_0, t, \omega)) - f(g(X_0, V_0, t, \omega))| dt \le B \ , \quad \text{and } \limsup_{\epsilon \downarrow 0} E^0_\epsilon(B) < \infty \, .$$

(A5)

$$E^0_1 \left[\int_0^\infty |f(g(X_0, W, t, \omega)) - f(g(X_0, V_0, t, \omega))| dt \right] < \infty \, ,$$

where, conditionally on $X_{T_{-1}}$, $W \overset{d}{=} \pi^{X_{T_{-1}}}$ and independent of $\mathcal{F}^X \vee \mathcal{F}^V \vee \mathcal{G}$.

We will show the following:

Theorem 1. *Under (A1)-(A5),*

$$E_\epsilon(f(Y_0)) = f^{(0)} + \epsilon f^{(1)} + o(\epsilon) \ , \ \text{as } \epsilon \downarrow 0 \, ,$$

where

$$f^{(0)} = E_1(f(Z_0)) = \sum_{x \in \mathcal{X}} P_1(X = x) \sum_{y \in \mathcal{Y}} \pi^x(y) f(y)$$

$$f^{(1)} = \lambda E^0_1 \left(\int_0^\infty [f(g(X, W, t, \omega)) - f(g(X, V, t, \omega))] dt \right) \, .$$

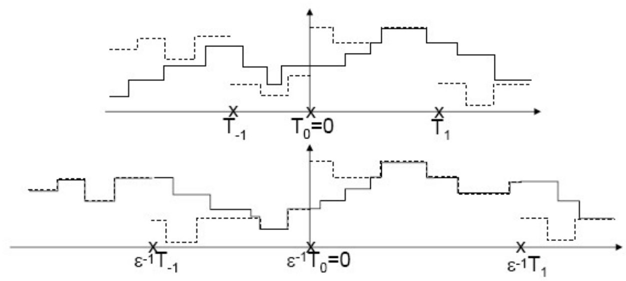

FIGURE 1: The argument behind the approximation in Theorem 1 based on comparison of sample paths of Y (solid) and Z (dashed line). In the lower part, the jump times of X are dilated by ϵ^{-1}.

Remark 1. The term $f^{(0)}$ can be computed in terms of the deviation matrix of a Markov chain [13], which is given as a solution to a matrix equation. In certain cases, most notably the M/M/1 queue, explicit formulas exist [12] (see the example following Theorem 2).

Remark 2. Functions $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ can be handled as well, without changing the proofs.

Remark 3. In the case that (X, Y) is a Markov chain, (A1) is equivalent to tightness of the time-stationary marginals. This fact is shown in appendix A.

Remark 4. The idea behind the approximation in Theorem 1 is the following. Assume that (X, Y) are ergodic processes even though this is not needed in the proof. Compute the "finite difference" $(E_\epsilon(f(Y_0)) - E_\epsilon(f(Z_0)))/\epsilon$ using the mean-cycle formula over the jumps of X. The difference is equal to the difference accumulated during a typical cycle, i.e.

$$\frac{E_\epsilon(f(Y_0)) - E_\epsilon(f(Z_0))}{\epsilon} = \lambda E_\epsilon^0 \left(\int_0^{T_1} [f(Y_t) - f(Z_t)]dt \right) , \tag{1}$$

This is a function of the sample paths of Y and Z between times T_0 and T_1. Now, consider dilations of the jump times of X by ϵ^{-1} as depicted in Figure 2. The assumption on tightness, (A1), ensures that Y_t and Z_t are not far apart at the beginning of each cycle. So in Figure 2 at T_{-1}, the difference of Y_t and Z_t does not grow as ϵ is decreased. If $\epsilon > 0$ is sufficiently small then Y_t and Z_t have ample time to couple by time 0, as it happens in the lower part of the figure. Indeed, (A2) will guarantee that coupling occurs during sufficiently long cycles. Hence Y_0 is approximately distributed as $\pi^{X_{0-}}$. Now, since $\epsilon^{-1}T_1 \to +\infty$ we expect the RHS of (1) to converge to $f^{(1)}$. This is valid as long as the limit can be exchanged with the expectation. This is justified by dominated convergence, which applies because of assumptions (A4) and (A5).

Proof. By (A3), the Palm inversion formula (see [9]) gives,

$$\frac{E_\epsilon(f(Y_0) - f(Z_0))}{\epsilon} = \lambda E_\epsilon^0 \left[\int_0^{T_1} (f(Y_t) - f(Z_t))dt \right] .$$

Since $E_\epsilon(f(Z_0)) = f^{(0)}$, to prove the theorem it suffices to show that the limit, as $\epsilon \downarrow 0$, of the RHS is $f^{(1)}$.

First we show $P_\epsilon^0[Y_0 \in \cdot | X_{0-}, X_0] \to \pi^{X_{0-}}(\cdot)$ as $\epsilon \downarrow 0$, P_1^0-a.s. . Since $P_\epsilon^0 \circ (X_{-0}, X_0)^{-1}$ does not depend on ϵ, tightness of $(P_\epsilon^0[(Y_0, T_1) \in \cdot | X_{-0} = x, X_0 = x'], \epsilon > 0)$ and $(P_\epsilon^0[(Y_0, T_1) \in \cdot | X_0 = x, X_{T_1} = x'], \epsilon > 0)$, for each x, x' with $P_\epsilon^0(X_0 = x, X_{T_1} = x') > 0$, follows. (These are distributions on $\mathbb{Z}_+ \times (\mathbb{R}_+ \cup \{\infty\})$, where $\mathbb{R}_+ \cup \{\infty\}$ is equipped with the Borel σ-field of the one-point compactification.) Since $\mathcal{X}$ is countable, by diagonalization we can extract a subsequence $\epsilon_k \downarrow 0$ such that $P_{\epsilon_k}^0[(Y_0, T_1) \in \cdot | X_{0-}, X_0] \to \nu^{X_{0-}, X_0}(\cdot) \times \delta_\infty$, as $k \to \infty$, P_1^0-a.s., where δ_∞ is the point-mass at ∞. We will show that $\nu^{X_{0-}, X_0} = \pi^{X_{0-}}$. Take any subsequence of (ϵ_k) (which we denote it again by (ϵ)) for which the limit $\mu^{X_0, X_{T_1}} \times \delta_\infty(\cdot)$ of $P_\epsilon^0[(Y_0, T_1) \in \cdot | X_0, X_{T_1}]$ converges P_1^0-a.s. . Now, take r.v's $((Y^\epsilon, T_1^\epsilon), \epsilon > 0)$ on P_1 (perhaps when enlarged), satisfying $P_1^0[(Y^\epsilon, T_1^\epsilon) \in \cdot | X_0, X_{T_1}] = P_\epsilon^0[(Y_0, T_1) \in \cdot | X_0, X_{T_1}]$ P_1^0-a.s., and whose P_1^0-a.s. limit (Y^0, ∞), as $\epsilon \downarrow 0$, is distributed according to $\mu^{X_0, X_{T_1}} \times \delta_\infty$. Moreover, $(Y^\epsilon, T_1^\epsilon)$ is taken to be conditionally independent of $\mathcal{F}^V \vee \mathcal{G}$ given X_0. Hence,

$$P_\epsilon^0[g(X_0, Y_0, T_1, \omega) \in \cdot | X_0, X_{T_1}] = P_1^0[g(X_0, Y^\epsilon, T_1^\epsilon, \omega) \in \cdot | X_0, X_{T_1}] . \tag{2}$$

Now, from (A2), $g(X_0, Y^0, T_1^\epsilon, \omega) = g(X_0, V_0, T_1^\epsilon, \omega)$, for small enough ϵ. Since $Y^\epsilon \to Y^0$, we have $g(X_0, Y^\epsilon, T_1^\epsilon, \omega) = g(X_0, Y^0, T_1^\epsilon, \omega)$ for small $\epsilon > 0$. Combining with (2) this yields,

$$P_\epsilon^0[g(X_0, Y_0, T_1, \omega) \in \cdot | X_0, X_{T_1}] - P_1^0[g(X_0, V_0, T_1^\epsilon, \omega) | X_0, X_{T_1}] \to 0 , \quad P_1^0\text{-a.s.} ,$$

so, the limit of the first term on the LHS is π^{X_0} by the conditional independence of V_0

and $\mathcal{G}$ given X_0. Notice that

$$P^0_\epsilon[g(X_0, Y_0, T_1, \omega) \in \cdot | X_0, X_{T_1}] \circ \theta^{T_1} = P^0_\epsilon[Y_0 \in \cdot | X_{0-}, X_0] \,,$$

so (2) implies $\nu^{X_{0-}, X_0} = \pi^{X_{0-}}$.

Redefine $(Y^\epsilon, T^\epsilon_1)$ such that it again has P^0_1-a.s. limit (Y^0, ∞), and $P^0_1[(Y^\epsilon, T^\epsilon_1) \in \cdot | X_{0-}, X_0] = P^0_\epsilon[(Y_0, T_1) \in \cdot | X_{0-}, X_0]$, and $(Y^\epsilon, T^\epsilon_1)$ is conditionally independent of $\mathcal{F}^V \vee \mathcal{G}$ given X_{0-}. Now by (A2),

$$\int_0^{T^\epsilon_1} [f(g(X_0, Y^\epsilon, t, \omega)) - f(g(X_0, V_0, t, \omega))]dt \underset{\epsilon \downarrow 0}{\longrightarrow} \int_0^\infty [f(g(X_0, Y^0, t, \omega)) - f(g(X_0, V_0, t, \omega))]dt < \infty \,, \quad P^0_1\text{-a.s.} \,.$$

(A4) implies that the sequence of r.v's on the LHS is uniformly integrable, so together with (A5) imply that the limit under E^0_1 exists and equals $f^{(1)}/\lambda$.

3. Application to queues

Here, we show that Jackson networks with varying arrival and service rates can be put in the framework of the previous section.

Let $J \geq 1$, and X be as in the previous section. On state x of the environment, $Y = (Y_t(i), i = 1, \ldots, J, t \in \mathbb{R})$ is a J-station open Jackson network with external arrival rates $(\lambda(x)R_{01}, \ldots, \lambda(X)R_{0J})^T$ for some $(R_{0i}, i = 1, \ldots, J) =: R_{0\cdot}$ with $R_{0i} \geq 0, \sum_{i=1}^J R_{0i} = 1$, service rates $\mu(x) := (\mu_1(x), \ldots, \mu_J(x))$, and a $J \times J$ routing matrix $R = (R_{ij}, i, j \in \{1, \ldots, J\})$. In other words, for any $f : \mathbb{Z}^J_+ \to \mathbb{R}$, the generator Q^x is given by

$$\begin{aligned} Q^x f(y) = &\sum_{i=1}^J \lambda(x) R_{0i}(f(y + e_i) - f(y)) \\ &+ \sum_{i=1}^J \sum_{j=1}^J \mu_i(x) 1(y_i > 0) R_{ij}(f(y - e_i + e_j) - f(y)) \\ &+ \sum_{i=1}^J 1(y_i > 0)\mu_i(x) R_{i0}(f(y - e_i) - f(y)) \,, \end{aligned} \tag{3}$$

where R is a stochastic matrix such that $I - R$ is non-singular, and e_i is the vector with zero elements except a 1 in the i-th coordinate. The vector of arrival rates at each node is given by $a(x) = \lambda(x)(I - R^T)^{-1} R_{0\cdot}$.

We now define the processes $g(x, y, t, \omega)$. For any $i \in \{1, \dots, J\}, j \in \{0, \dots, J\}$, let N_{0i}^x, N_{ij}^x be Poisson processes with intensities $\lambda(x)R_{0i}$ and $\mu_i(x)R_{ij}$, respectively, such that $(N_{0i}^x, N_{ij}^x | i \in \{1, \dots, J\}, j \in \{0, \dots, J\})$ are independent. Define $g(x, y, t, \omega)$ to be the solution of the SDE

$$U_i(t) = y_i + N_{0i}^x(t) + \sum_{j=1}^J \int_0^t 1\{U_j(s-) > 0\} N_{ji}^x(ds) - \sum_{j=0}^J \int_0^t 1\{U_i(s-) > 0\} N_{ij}^x(ds) \, ,$$

$i = 1, \dots, J$, w.r.t. $(U(t), t \geq 0)$.

We will consider the assumptions,

(B1) $\tilde{\rho}_i := \sup_x a_i(x)/\mu_i(x) < 1$ for all i,

(B2) $\inf_x \mu_i(x) > 0$ for all i, and $\inf_x \lambda(x) > 0$,

and show the following

Theorem 2. *Suppose (B1)-(B2) hold. Then,*

(a) (X, Y) possesses a stationary law, for each $\epsilon > 0$.

(b) (A1)-(A5) are satisfied, and therefore $P_\epsilon(Y_0 = y)$ can be approximated using Theorem 1.

(Notice that in part (b), we invoke Theorem 1 for $f(Y_0) = 1\{Y_0 = y\}$, where $y \in \mathcal{Y}$.)

Remark 5. Assumption (B2) regarding the service rates can be weakened by putting restrictions on X. What is necessary is that equations (4)-(6) below, possess a unique solution.

Example.[M/M/1 queue] Here, $J = 1$ and $\pi^x(y) = \rho(x)(1 - \rho(x))^y$. Using the formula in [12] for the deviation matrix of an M/M/1 queue, we have

$$P_\epsilon(Y_0 = y) = \sum_x P_1(X_0 = x)\pi^x(y) + \epsilon\lambda \sum_x P_1^0(X_{0-} = x, X_0 = x') \cdot \sum_{i=1}^\infty \pi^x(i) \frac{\rho(x')^{(y-i)^+} - (i + y - 1)\pi^{x'}(y)}{\mu(x') - \lambda(x')} + o(\epsilon) \, , \quad \text{as } \epsilon \downarrow 0 \, .$$

In sections 3.2 and 3.3, parts (i) and (ii) are proved, respectively. Specifically, in section 3.2 we construct a stationary version of (X, Y). To do this we compare Y with the queue lengths $\tilde{Y}$ of another Jackson network with service speeds $\tilde{\mu}(x) \leq \mu(x)$ for all

x. $\tilde{Y}$ is constructed by applying a time-change to a non-varying Jackson network. The existence of a stationary version for the latter is established in [3]. In section 3.3, again key is the comparison with $\tilde{Y}$. The fact that the marginal distribution of $\tilde{Y}$ does not depend on ϵ and the monotonicity of $g(x, y, t, \omega)$ in $y \in \mathcal{Y}$, reduce the proofs of (A4) and (A5) into showing that the mean residual busy period of some non-varying Jackson network is finite. This follows from the fact that in Jackson networks, recurrence times possess finite moments of all orders (see [14] and Theorem 15.0.1 in [15]).

In the next section we review the concept of an Euler network from Baccelli and Foss [3] modified to allow for variable service speeds. This is used for establishing the monotonicity properties in sections 3.2 and 3.3.

3.1. Preliminaries

We follow the construction in [3] allowing for variable service speeds. A tuple $\Sigma = (N, t, \sigma, r, \mu)$ is an *Euler network* (see [3]) when

- $t = (t(1), \ldots, t(N)), N \geq 1$ is a sequence of non-decreasing real numbers.
- $\sigma = (\sigma_i(n) \in \mathbb{R}_+,\ i = 1, \ldots, J, n = 0, \ldots, d_i)$, for some $d_i \geq 0$. We call σ, the *service sequence*, and $\sigma_i(n)$ is interpreted as the job size of the n-th customer served by station $i \geq 1$.
- $r = (r_i(n) \in \{0, \ldots, J\},\ i = 0, \ldots, J, n = 0, \ldots, d_i)$ for the d_i's above. r is called the *switching sequence*, and $r_i(n)$ is interpreted as the station that the n-th serviced customer from station i jumps to.
- $\mu = (\mu_i(t) \in \mathbb{R}_+,\ i \in \{1, \ldots, J\}, t \geq 0)$ with $\inf_t \mu_i(t) > 0$. $\mu_i(t)$ is interpreted as the instantaneous server speed at the i-th station.
- Under the above interpretations of t, σ, r, μ, the number of arrivals and departures from station i both equal d_i.

By Theorem 8 in [3], for such Σ, the number of arrivals and departures from station i is determined by r, N alone.

Given Σ, the set of departure times $D = (D_j(n), j = 1, \ldots, J, n = 1, \ldots, d_j)$ is determined by the equations

$$D_0(n) = t_n \, , \quad 1 \leq n \leq N \, , \tag{4}$$

$$\int_{S_j(n)}^{D_j(n)} \mu_j(t) ds = \sigma_j(n) \, , \quad j = 1, \ldots, J \, , \tag{5}$$

$$S_j(n) := \max(D_j(n-1), \min_{n_0 + \ldots + n_k = n} (\max_{i=0,\ldots,J} D_i(\eta_{i,j}(n_i))) \, , \quad j = 0, \ldots, J \, , \tag{6}$$

where

$$\eta_{i,j}(n) = \inf\{d_i \geq m \geq 1 | \sum_{p=1}^{m} 1\{r_i(p) = j\} = n\} \, , \quad i,j = 0, \ldots, J \, ,$$

is the minimum number of service completions at station i such that n of them are routed to j. The assumption $\inf_t \mu_i(t) > 0$ is made so that (4)-(6) always have a unique solution. (For more on evaluating (4)-(6) inductively, see [10].) We will sometimes write $d_j(\Sigma), D(\Sigma) = (D_j(\Sigma, n), j = 1, \ldots, J, n = 1, \ldots, d_j(\Sigma)), S_j(\Sigma, n)$ to emphasize the dependence on Σ.

Define the time to empty after the last arrival $\max_{i=1,\ldots,J} \max_{n=1,\ldots,d_i} D_i(n) - t_N =: Z(\Sigma)$, and the size of queue j at time s by

$$Q_j(\Sigma, s) = \sum_{i=0}^{J} \sum_{n=1}^{d_i} 1\{D_i(n) < s \, , \, r_i(n) = j\} - \sum_{n=1}^{d_j} 1\{D_j(n) < s\} \, .$$

Lemma 1. *Let $\Sigma^k = (N, t^k, \sigma, r, \mu^k), k = 1, 2$ be two Euler networks with $\mu^1 \geq \mu^2$, $t^1 \leq t^2$. Then, $D(\Sigma^1) \leq D(\Sigma^2)$.*

Proof. Follows directly by applying induction to (4)-(6).

Before obtaining the important monotonicity property of Lemma 2, we need the concept of composition from [3]. The composition $\Sigma := (N, t, \sigma, r, \mu) = \Sigma^1 + \Sigma^2$, where

$\Sigma^k = (N^k, t^k, \sigma^k, r^k, \mu), k = 1, 2$ satisfying $t^1(N^1) \le t^2(1)$, is defined as follows:

$$
\begin{aligned}
N &= N^1 + N^2 \\
t &= (t^1(1), \dots, t^1(N^1), t^2(1), \dots, t^2(N^2)) \\
r_i(n) &= \begin{cases} r_i^1(n) & 1 \le n \le d_i(\Sigma^1) \\ r_i^2(n - d_i(\Sigma^1)) & d_i(\Sigma^1) < i \le d_i(\Sigma^1) + d_i(\Sigma^2) \end{cases} \\
\sigma_i(n) &= \begin{cases} \sigma_i^1(n) & 1 \le n \le d_i(\Sigma^1) \\ \sigma_i^2(n - d_i(\Sigma^1)) & d_i(\Sigma^1) < i \le d_i(\Sigma^1) + d_i(\Sigma^2) \end{cases}
\end{aligned}
$$

By Theorem 9 in [3], $d_i(\Sigma) = d_i(\Sigma^1) + d_i(\Sigma^2)$ which follows from the fact that the number of arrival and departures equals $d_i(\Sigma^\cdot)$, regardless of the timing information in t, σ. This also implies that composition is associative.

Lemma 2. $Z(\Sigma^1 + \Sigma^2) \ge Z(\Sigma^2)$.

Proof. Define $\Sigma^1(\Delta) = (N^1, (t^1(n) - \Delta, n = 1, \dots, N^1), \sigma^1, r^1, \mu)$, for $\Delta > 0$. By (5),

$$D_i(\Sigma^1(\Delta), n) \le \frac{\sigma_i^1(n)}{\inf_t \mu_i(t)} + S_i(\Sigma^1(\Delta), n) \, ,$$

for n, i such that all quantities are defined. Since no more than $\sum_{k=1}^J d_k(\Sigma^1)$ induction steps of (5)-(6) are needed in evaluating $S_i(\Sigma^1(\Delta), n)$, we get

$$D_i(\Sigma^1(\Delta), n) - (t^1(N^1) - \Delta) \le \sum_{k=1}^J d_k(\Sigma^1) \max_{j=1,\dots,J} \frac{\sigma_j^1(n)}{\inf_t \mu_j(t)} \, ,$$

with the LHS not depending on Δ. Since $\inf_t \mu_i(t) > 0$ for all i, we can have $Z(\Sigma^1(\Delta)) + t^1(N^1) - \Delta < t^2(1)$ for some large enough Δ. For such Δ, $Z(\Sigma^1(\Delta) + \Sigma^2) = Z(\Sigma^2)$, so by Lemma 1, $Z(\Sigma^2) \le Z(\Sigma^1 + \Sigma^2)$

We will need one more definition; that of a time-changed Euler network. Let $\tau : \mathbb{R} \to \mathbb{R}$ be a strictly increasing bijection. For an Euler network $\Sigma = (N, t, \sigma, r, \mu)$, define $\tau\Sigma := (N, \tau(t), \sigma, r, \nu)$, where $\nu(s) = \mu(\tau^{-1}(s))/\tau'(\tau^{-1}(s))$ for all $s \in \mathbb{R}$ and $\tau'(\cdot)$ is the derivative of $\tau(\cdot)$. (Notice that ν is defined Lebesgue-a.e. .) It is immediate from (4)-(6) that the following holds.

Lemma 3. $D(\tau\Sigma) = \tau(D(\Sigma))$, $Q(\Sigma, s) = Q(\tau\Sigma, \tau(s))$ *for all* $s \in \mathbb{R}$.

3.2. Proof of Theorem 2(a)

Let $(t(k), k \in \mathbb{Z})$ be a Poisson process with stochastic intensity $\lambda(X_t)$ conditional on $\mathcal{F}^X$. For each $k \in \mathbb{Z}$, let r^k be a switching sequence generated by a random path in the Jackson network, with probabilities given by the routing matrix (R_{ij}). Moreover, for each such r^k, let σ^k be the associated service sequence, such that the $\sigma_j^k(n)$'s are i.i.d. exponential r.v. with mean one, and independent of everything else. Let $\Sigma^k := (1, \{t(k)\}, \sigma^k, r^k, \mu(X.))$, $\tilde{\Sigma}^k := (1, \{t(k)\}, \sigma^k, r^k, \tilde{\mu}(X.))$ for $k \in \mathbb{Z}$. For $k, m \in \mathbb{Z} \cup \{-\infty, +\infty\}$, with $m \geq k$, we write $\Sigma_k^m := \Sigma^k + \ldots + \Sigma^m$ and similarly for $\tilde{\Sigma}_k^m$. Also, consider the time-change

$$\mathbb{R} \ni s \mapsto \tau(s) = \inf \left\{ v \in \mathbb{R} | s = \int_0^v 1/\lambda(X_u) du \right\} .$$

By (B2), this is bijective, so Lemma 3 can be applied.

Since $\tilde{\mu}(X_s) \leq \mu(X_s)$ for all $s \in \mathbb{R}$, by Lemma 1 we have $D(\tilde{\Sigma}_k^0) \geq D(\Sigma_k^0)$ for all $k \leq 0$. This and Lemma 2 yields $Z(\tilde{\Sigma}_{-\infty}^0) \geq Z(\Sigma_{-\infty}^0)$. Now, let $N_s, s \in \mathbb{R}$ be the counting process of points $(t(k), k \in \mathbb{Z})$, (i.e., $N_s = \sum_{k \in \mathbb{Z}} 1\{s \leq t(k) \leq 0\}$, for $s \leq 0$). Then, say if $s_2, s_1 \in \mathbb{R}$, $s_2 \leq s_1 \leq 0$,

$$E[N_{\tau^{-1}(s_2)} | \mathcal{F}^N_{\tau^{-1}(s_1)} \vee \mathcal{F}^X] = E\left[\int_{s_1}^{\tau^{-1}(s_2)} \lambda(X_u) dt \middle| \mathcal{F}^N_{\tau^{-1}(s_1)} \vee \mathcal{F}^X \right] ,$$

where $\mathcal{F}_u^N$ is the filtration generated by histories of N from time $u \leq 0$ to $+\infty$. But by the definition of τ,

$$\int_{s_1}^{\tau^{-1}(s_2)} \lambda(X_u) du = s_2 - s_1 , \quad \forall s_1, s_2, s_2 \leq s_1 \leq 0 .$$

Therefore, $(\tau(t(k)), k \in \mathbb{Z})$ are the jump times of a Poisson process with unit intensity. Notice also, that $\tilde{\mu}_i(\tau^{-1}(s))/\tau'(\tau^{-1}(s)) = \tilde{\rho}_i^{-1} e_i^T (I - R^T)^{-1} R_{0\cdot}^T$. Hence, $\tau\tilde{\Sigma}$ corresponds to a Jackson network with constant arrival and service rates. Notice that the arrival rate $e_i^T (I - R^T)^{-1} R_{0\cdot}^T$ at queue i, is less than the service rate, so $Z(\tau\tilde{\Sigma}_{-\infty}^m) \leq \tau(t(m+1)) - \tau(t(m))$ for some $m < 0$ (see [3]). By Lemmas 3 and 2, this implies $Z(\tilde{\Sigma}_{-\infty}^m) \leq t(m+1) - t(m)$, i.e., the network empties before the $m+1$-st arrival. Therefore, $Z(\Sigma_{-\infty}^m) \leq t(m+1) - t(m)$ and $Q(\Sigma_{-\infty}^0, 0) := Q(\Sigma_{m+1}^0, 0)$ is well-defined. Now, $Y(t) := Q(\Sigma_{-\infty}^0, 0) \circ \theta^t, t \in \mathbb{R}$ defines a (P, θ)-compatible process, for which it is straightforward to show that conditionally on $\mathcal{F}^X$, it is a Markov chain with generator given by (3) on $\{X_t = x\}$.

3.3. Proof of Theorem 2(b)

Let m be as in the proof of part (a). Since $D(\Sigma^0_{m+1}) \le D(\tilde\Sigma^0_{m+1})$, we have

$$\sum_{i=1}^J Y_0(i) \le \sum_{i=1}^J Q_i(\tilde\Sigma^0_{m+1}, 0)\,. \tag{7}$$

But $Q_i(\tilde\Sigma^0_{m+1}, 0) = Q_i(\tau\tilde\Sigma^0_{m+1}, 0)$ is a geometric r.v. with mean $1/(1-\tilde\rho_i)$ as the invariant queue length in a Jackson network. This implies that $(P^0_\epsilon(Y_0(i) \in \cdot), \epsilon > 0)$ is tight. Also (A3) is satisfied.

Let $y^2 := (y^2_1, \ldots, y^2_J) \ge (y^1_1, \ldots, y^1_J) =: y^1$ be vectors with nonnegative components, and consider the processes $t \mapsto (g(x, y^k, t, \omega), k = 1, 2)$ defined in the beginning of this section. Say $g(x, y^2, s, \omega) \ge g(x, y^1, s, \omega)$ for all $0 \le s < t$, and at t one of the two processes jumps. Then, if $g_j(x, y^2, t, \omega)$ jumps downwards then so does $g_j(x, y^1, t, \omega)$ provided it is nonzero. If $g_j(x, y^1, t, \omega)$ jumps upwards then so does $g_j(x, y^2, t, \omega)$ since, in the case of a non-external arrival the upstream queue i will have $g_i(x, y^2, t-, \omega) \ge g_i(x, y^1, t-, \omega) > 0$. Hence, $g(x, y^2, t, \omega) \ge g(x, y^1, t, \omega)$ for all $t \ge 0$. This implies that (A2) holds.

It remains to show (A4), (A5). By the monotonicity above,

$$\begin{aligned}
&\int_0^\infty |1\{g(X_0, Y_0, t, \omega) = y\} - 1\{g(X_0, V_0, t, \omega) = y\}| dt \\
&\qquad\qquad \le \inf\{t > 0 | g(X_0, Y_0, t, \omega) = g(X_0, V_0, t, \omega) = 0\} \\
&\qquad\qquad \le \inf\{t > 0 | g(X_0, Y_0 \vee V_0, t, \omega) = 0\}\,.
\end{aligned}$$

By (7), and the definition of V_0, $\sum_i (Y_0(i) \vee V_0(i)) \le \sum_{i=1}^J Q^k_i$, $k = 1, 2$ for some i.i.d. Q^1, Q^2 each distributed as the RHS in (7). Thus, the integral above is bounded by

$$\inf\left\{ t > 0 | g(X_0, \left(\sum_{i=1}^J Q^1_i \vee Q^2_i\right) \sum_{j=1}^J e_j, t, \omega) = 0 \right\} =: B\,.$$

Now, on the same probability space we can define independent geometric r.v's $(Z_i, i = 1, \ldots, J)$, each with mean $1/(1-\hat\rho_i) > 0$, such that

$$\sum_i Q^1_i \vee Q^2_i \le \wedge_{i=1}^J Z_i\,, \quad P^0_1\text{-a.s.} \tag{8}$$

on a set $\{\sum_i Q^1_i \vee Q^2_i > M\}$, where M is a finite constant. This is because the LHS of (8) has a geometric tail which can be dominated by the tail of the RHS for $\hat\rho_1, \ldots, \hat\rho_J$

close enough to 1. Thus,

$$E_1^0(B) \le E_1^0(\inf\{t > 0 | g(X_0, M\sum_{j=1}^{J} e_j, t, \omega) = 0\})$$
$$+ E_1^0(\inf\{t > 0 | g(X_0, (Z_1, \ldots, Z_J)^T, t, \omega) = 0\})\,.$$

To show that this is finite, it suffices to do it for the second term which we compare with the time to empty for an alternative Jackson network, when started with the invariant distribution.

Let $\hat{\mu}_i(x) = a_i(x)/\hat{\rho}_i$ for all i, and $\hat{\Sigma}(k) = (1, t(k), \sigma^k, r^k, \hat{\mu}), k \in \mathbb{Z}$. Now $\tau\hat{\Sigma}$ corresponds to a Jackson network with unit arrival rate and service rate $\hat{\rho}_i e_i^T (I - R^T)^{-1} R_{0\cdot}^T$ at station i, so $Q(\hat{\Sigma}_{-\infty}^0, 0) = Q(\tau\hat{\Sigma}_{-\infty}^0, 0) \overset{d}{=} (Z_1, \ldots, Z_J)^T$. Let $n(t) = \max\{m \in \mathbb{Z} | m \le t\}$. Now,

$$\begin{aligned}
&E_1^0(\inf\{t > 0 | g(X_0, (Z_1, \ldots, Z_J)^T, t, \omega) = 0\}) \\
&\qquad = E_1^0(\inf\{t > 0 | Q(\hat{\Sigma}_{-\infty}^0 + \Sigma_1^{n(t)}) = 0\}) \\
&\qquad \le E_1^0(\inf\{t > 0 | Q(\hat{\Sigma}_{-\infty}^{n(t)}, t) = 0\}) \\
&\qquad \le \frac{1}{\inf_x \lambda(x)} E_1^0(\tau(\inf\{t > 0 | Q(\hat{\Sigma}_{-\infty}^{n(t)}, t) = 0\})) \\
&\qquad \le \frac{1}{\inf_x \lambda(x)} E_1^0(\inf\{t > 0 | Q(\tau\hat{\Sigma}_{-\infty}^{n(t)}, t) = 0\})) < \infty\,,
\end{aligned}$$

since this is the mean time to hit the empty state, starting from the invariant distribution. This shows that (A4) is satified. Since $W \le_{\text{st}} Q^1$, where W is as in Theorem 1, the same arguments show that (A5) holds as well.

Appendix A.

We show the following statement: if X is a Markov chain then (A1) is equivalent to tightness of $(P_\epsilon(Y_0 \in \cdot), \epsilon > 0)$.

Proof. Let $q(X_t)$ be the instantaneous rate of jumps for X at time t. Since the marginal distribution of X does not depend on the parameter ϵ, if $(P_\epsilon(Y_0 \in \cdot), \epsilon > 0)$ is tight, it follows that $(P_\epsilon[Y_0 \in \cdot | X_0 = x], \epsilon > 0)$ is tight.

For fixed $\alpha > 0$ and any x, let $F_x \subset \mathcal{X}$ be finite and such that $\inf_{\epsilon > 0} P_\epsilon[Y_0 \in$

$F_x | X_0 = x] > 1 - \alpha$. Now, for fixed arbitrary $\beta > 0$ let D be a finite set in $\mathcal{X}$ such that

$$E_\epsilon[q(X_0)1\{X_0 \in D\}] > (1-\beta)E_\epsilon(q(X_0)) \,, \quad \text{for all } \epsilon > 0$$

(because of the assumption that the marginal distribution of X does not depend on $\epsilon > 0$), and define $F := \cup_{x \in D} F_x$. Then by Papangelou's formula (see [9]),

$$\begin{aligned} P^0_\epsilon(Y_0 \in F) &= \frac{E_\epsilon[1\{Y_0 \in F\} q(X_{0-})]}{E_\epsilon(q(X_{0-}))} \\ &\geq \frac{E_\epsilon[P_\epsilon[Y_0 \in F | X_{0-}] q(X_{0-}) 1\{X_{0-} \in D\}]}{E_\epsilon(q(X_0))} \\ &\geq (1-\alpha)(1-\beta) \,, \quad \text{for any } \epsilon > 0 \,, \end{aligned}$$

where we have used the observation that X, Y have no simultaneous jumps (P_ϵ and P^0_ϵ a.s.).

Since α, β were arbitrary, tightness of the LHS follows.

References

[1] Loynes, R.M. (1962.) The stability of a queue with non independent inter-arrival and service times, *Proc. Camb. Philos. Soc.*, **58**, 497–520.

[2] Neuts, M. (1978.) The M/M/1 queue with randomly varying arrival and service rates, *Opsearch*, **15**, 139–157.

[3] Baccelli, F. and Foss, S. (1994). Ergodicity of Jackson-Type Queueing Networks, *Queueing Systems*, **17**, 5–72.

[4] Gelenbe, E. and Rosenberg, C. (1990). Queues with Slowly Varying Arrival and Service Processes, *Management Science*, **36**, 928–937.

[5] Baccelli, F. and Brémaud, P. (2003). Virtual Customers in Sensitivity and Light Traffic Analysis via Campbell's Formula for Point Processes, *Adv. Appl. Prob.*, **25**, 221–234.

[6] Blaszczyszyn, B. (1995). Factorial moment expansion for stochastic systems, *Stoch. Proc. Appl.*, **56**, 321–335.

[7] Baccelli, F. and Makowski, A. (1986). Queues in Random Environment: Stability and Bounds, *Stochastic models*, **2**, 281–291.

[8] Altman, E., Avrachenkov, K.E. and Núñez-Queija, R. (2004). Perturbation analysis for denumberable Markov chains with application to queueing models, *Adv. Appl. Prob.*, **36**, 839–853.

[9] Baccelli, F. and Brémaud, P. (2003). *Elements of Queueing Theory*, 2nd edn, Springer-Verlag.

[10] Baccelli, F., Cohen, G. and Gaujal, B. (1992) Recursive equation and basic properties of timed Petri nets, *Journal of Dynamical Discrete Event Systems*, **1**, 415–438.

[11] Yin, G.G. and Zhang, Q. (1998). *Continuous-time Markov Chains and Applications*, Springer-Verlag, New York, 1998.

[12] Koole, G. (1998). The deviation matrix of the M/M/1 and M/M/1/N queue, with applications to controlled queueing models, *in Proc. of the 37th IEEE CDC*, Tampa, 56–59.

[13] Coolen-Schrijner, P. and van Doorn, E.A. (2002). The deviation matrix of a continuous-time Markov chain, *Probab. Engrg. Inform. Sci.*, **16**, 351–366.

[14] Fayolle, G., Malyshev, V.A., Menshikov, M.V. and Sidorenko, A.F. (1993). Lyapunov functions for Jackson networks, *Mathematics of Operations Research*, **18**, 916–927.

[15] Meyn, S.P. and Tweedie, R.L. (1993). Markov Chains and Stochastic Stability, Springer, London, 1993.