Non-ergodicity of queueing networks under non-stability of their fluid models

A. Puhalskii *
University of Colorado, Denver
and
Institute for Problems in Information Transmission, Moscow
A. Rybko †
Institute for Problems in Information Transmission, Moscow

November 15, 1998

Abstract

We study ergodicity properties of open queueing networks, for which the associated fluid models have trajectories that go to infinity. It is proved that if a trajectory is stable in a certain sense and tends to infinity linearly fast, then the underlying stochastic process is non-ergodic. The result applies to the basic non-trivial examples of non-ergodic networks found by Bramson, and Rybko and Stolyar. The proof employs some general large-deviation-theory results.

1 Introduction

In [16] Rybko and Stolyar observed that the condition that the load at every node of a multiclass queueing network is less than 1 is not sufficient for the network to be ergodic. In connection with this, they introduced a new approach to the analysis of ergodicity of networks, which reduces the problem to studying stability of the associated fluid models. It was shown that a certain two-node priority network is ergodic if and only if for every initial state of its fluid model the "total amount of fluid" eventually vanishes. This approach was further developed by Dai [5] and Stolyar [18], who proved that stability of the fluid model is sufficient for ergodicity of a general network. Interesting instances of non-ergodic queueing networks with loads at the nodes less than 1 were considered by Bramson [2, 3]. Historically, the first example of a network with loads less than 1, which has an unboundedly growing queue, was given by Kumar and Seidman [8]. Our purpose here is to investigate necessity of the Rybko-Stolyar condition, i.e., find out when non-stability of the fluid model implies that a queueing network is non-ergodic.

We consider an open multiclass queueing network of M nodes and L customer classes. There are $L' \leq L$ exogenous arrival processes, each consisting of customers of the same class. A customer's class uniquely specifies the node where the customer is served, so $L \geq M$. At a given node, customers form a single queue in the order of their arrivals and keep their positions in the queue until service is completed. They are picked for service according to the node's service discipline,

*Partially supported by grant RFBR 98-01-00062

†Partially supported by Swiss National Science Foundation, project 7SUPJ048214, INTAS-93-820 and RFBR 96-01-10020

which is assumed to take into account only the customers' classes and their positions in the queue; within a class, customers are served according to the FIFO discpline. On completing service a customer either leaves the network or changes its class and transfers to the corresponding node.

We associate with the network a fluid model and show that under certain technical conditions the following holds: if there exists an initial state of the fluid model such that uniformly over all "fluid trajectories" with close initial states (note that in general the fluid trajectories are non-unique) the total amount of fluid grows to infinity at least linearly fast, then the random process that characterises the network is non-ergodic. This result complements the results of Dai [5] and Stolyar [18] and apparently covers all existing examples of non-ergodic networks. It also appears to be easier to apply than the result of Meyn [12], who, considering a Markovian setting, proved that if all "fluid-limit trajectories" tend to infinity, then the underlying stochastic process is non-recurrent. As for the proof of the theorem, a critical element is the use of large deviation theory, Freidlin and Wentzell [4], Varadhan [19], Puhalskii [13], to obtain estimates of the rates of convergence in the laws of large numbers that give rise to fluid limits. This general approach allows us to prove the results under fairly general assumptions on arrival, service and routing processes, in particular, we do not require that the associated random variables be i.i.d.

The paper is organised as follows: Section 2 contains technical preliminaries; in Section 3 we introduce stochastic processes characterising the network, define an associated fluid model and state the main result, which is proved in Section 4; in Section 5 we discuss the conditions of the theorem, properties of "fluid dynamics" and consider examples.

2 Technical preliminaries.

We first recall some facts from large deviation theory, see Freidlin and Wentzell [4], Varadhan [19]. Let U be a metric space with Borel σ-algebra $\mathcal{B}(U)$. The function $I : U \to [0,\infty]$ is said to be a rate function if the sets $\{z \in U : I(z) \le c\}$ are compact for all $c \ge 0$. Below we often use the simple fact that rate functions attain infima on closed sets. A sequence $\{P_n, n \ge 1\}$ of probability measures on $(U, \mathcal{B}(U))$ (or a sequence $\{X_n, n \ge 1\}$ of random elements with distributions P_n) is said to obey a large deviation principle (LDP) for a normalising sequence $a_n \to \infty$ with rate function I, if

$$\underline{\lim}_{n\to\infty} \frac{1}{a_n} \log P_n(G) \ge -\inf_{z\in G} I(z)$$

for all open sets G and

$$\overline{\lim}_{n\to\infty} \frac{1}{a_n} \log P_n(F) \le -\inf_{z\in F} I(z)$$

for all closed sets F.

Lemma 2.1 *If a sequence $\{P_n, n \ge 1\}$ obeys an LDP for a normalising sequence a_n with rate function I, then for every sequence of Borel sets A_n*

$$\overline{\lim}_{n\to\infty} \frac{1}{a_n} \log P_n(A_n) \le -\inf_{z\in A} I(z),$$

where $A = \bigcap_{l=1}^{\infty} cl \left(\bigcup_{n=l}^{\infty} A_n\right)$, and cl denotes taking the closure of a set.

The proof applies the argument of the proof of the upper bound in Puhalskii [14, Theorem 2.1] to the functions $f_n(z) = 1(z \in A_n)$.

Let f be a Borel function from a metric space U into a metric space U'. According to the contraction principle, Varadhan [19], Freidlin and Wentzell [4], if a sequence $\{P_n, n \ge 1\}$ of probability measures on $(U,\mathcal{B}(U))$ obeys an LDP for a normalising sequence a_n with rate function I and if the mapping f is continuous, then the sequence $\{P_n\circ f^{-1}, n\ge 1\}$ of images of P_n under f obeys an LDP (in U') for the normalising sequence a_n with rate function $I'(z')=\inf_{z\in f^{-1}(z')} I(z)$, $z'\in U'$. The extended contraction principle, Puhalskii [13], Puhalskii and Whitt [15], states that the latter LDP also holds if f is only continuous at points z such that $I(z)<\infty$. The following definition is prompted by the extended contraction principle: function $f:U\to U'$ is said to be U_0–continuous, where $U_0\subset U$, if f is continuous at every point of U_0, Puhalskii [14].

The next lemma extends the well-known fact that an LDP implies a "local" LDP, see, e.g., Freidlin and Wentzell [4], the proof being analogous.

Lemma 2.2 *Let $r(z,z')$ be a nonnegative function on $U\times U$, which, when restricted to $U_0\times U_0$, where $U_0\subset U$, is a metric on U_0 compatible with the topology induced on U_0 by the metric of U. Let sequence $\{P_n, n\ge 1\}$ obey an LDP in U for a normalising sequence a_n with rate function I. If $I(z)=\infty$ when $z\in U\setminus U_0$ and the function $r(z,z')$ is U_0–continuous in z' for all $z\in U_0$, then, for $z\in U_0$,*

$$I(z)=-\lim_{\varepsilon\to 0}\ \overline{\lim_{n\to\infty}}\ \frac{1}{a_n}\log P_n\left(z':\ r(z,z')\le\varepsilon\right)=-\lim_{\varepsilon\to 0}\ \underline{\lim}_{n\to\infty}\ \frac{1}{a_n}\log P_n\left(z':\ r(z,z')\le\varepsilon\right).$$

We next say that sequence $\{P_n, n\ge 1\}$ (or the associated sequence $\{X_n, n\ge 1\}$) is exponentially tight of order a_n if for every $\varepsilon>0$ there exists a compact $K\subset U$ such that $\overline{\lim}_{n\to\infty}\ P_n^{1/a_n}(U\setminus K)\le\varepsilon$, Deuschel and Stroock [6]. Recall, Puhalskii [13], that if a sequence $\{P_n, n\ge 1\}$ is exponentially tight of order a_n, then there exists a subsequence n' such that the sequence $\{P_{n'}, n'\ge 1\}$ obeys an LDP for the normalising sequence $a_{n'}$ with a rate function I'. Every such rate function I' is referred to below as a large-deviation (LD) accumulation point of $\{P_n, n\ge 1\}$ for the normalising sequence a_n. Also, we say that the sequence $\{P_n, n\ge 1\}$ is U_0-exponentially tight of order a_n, where $U_0\subset U$, if $\{P_n, n\ge 1\}$ is exponentially tight of order a_n and every its LD accumulation point I' for the normalising sequence a_n is such that $I'(z)=\infty$ for all $z\in U\setminus U_0$.

Let us introduce now a number of topological spaces, which we need below. Let S be a complete separable metric space. We denote by $D([0,\infty),S)$ the Skorohod space of right-continuous with left limits functions from $[0,\infty)$ to S equipped with the Skorohod–Prohorov–Lindvall metric, under which it is a complete separable metric space (see Skorohod [17], Billingsley [1], Lindvall [9], Liptser and Shiryaev [10], and Ethier and Kurtz [7] for more detail).

We denote by $\mathcal{V}_b^+$ the space of nondecreasing and bounded right-continuous functions on $[0,\infty)$ with values in $\mathbf{R}_+$, which are equal to 0 at 0. The following construction, borrowed from Liptser and Shiryaev [10], turns $\mathcal{V}_b^+$ into a complete separable metric space. Let $D([0,1],\mathbf{R})$ be the Skorohod space of real-valued right-continuous with left limits functions on $[0,1]$, and $D_0^+([0,1],\mathbf{R})$ be the subspace of $D([0,1],\mathbf{R})$ of nondecreasing functions, which are equal to 0 at 0 and left continuous at 1. We establish a one-to-one correspondence between $\mathcal{V}_b^+$ and $D_0^+([0,1],\mathbf{R})$ by mapping a function $f\in\mathcal{V}_b^+$ into the function $\hat f=(\hat f(t), t\in[0,1])$ defined by $\hat f(t)=f(-\log(1-t)), 0\le t<1$, and $\hat f(1)=\lim_{t\to\infty} f(t)$. Denoting by d_0 the (complete) Skorohod–Prohorov metric on $D([0,1],\mathbf{R})$, Billingsley [1], we set by definition $\rho(f,g)=d_0(\hat f,\hat g)$ for $f,g\in\mathcal{V}_b^+$. It is clear that $\mathcal{V}_b^+$ endowed with metric ρ is a metric space, which is homeomorphic to a closed subset of $D([0,1],\mathbf{R})$, so, it is a complete separable metric space. For $f=(f_j, j=1,\dots,J)\in(\mathcal{V}_b^+)^J$, we denote $||f||=\lim_{x\to\infty}\sum_{j=1}^J f_j(x)$.

Let also $D^+([0,\infty),\mathbf{R}^d)$ be the subspace of $D([0,\infty),\mathbf{R}^d)$ of componentwise nondecreasing and equal to 0 at 0 functions, $C([0,\infty),S)$ (respectively, $C^+([0,\infty),\mathbf{R}^d)$) be the subspace of continuous functions from $D([0,\infty),S)$ (respectively, $D^+([0,\infty),\mathbf{R}^d)$), $\mathcal{V}_{c,1}^+$ be the subspace of $D([0,\infty),\mathbf{R})$ of nondecreasing Lipshitz-continuous with Lipshitz constant 1 functions, equal to 0 at 0, and $\mathcal{V}^+_{b,c,1}$ be the subspace of $\mathcal{V}^+_b$ defined as $\mathcal{V}^+_{b,c,1} = \mathcal{V}^+_b \bigcap \mathcal{V}^+_{c,1}$. All the subspaces are endowed with relative topologies. Product spaces are endowed with product topologies.

Lemma 2.3 *Let $X^n = (X^n(t), t \ge 0), n \ge 1$, be stochastic processes on a probability space $(\Omega, \mathcal{F}, P)$ with paths from $D([0,\infty), S)$, where S is a complete separable metric space with metric d. Let S_0 be a subset of S. The sequence $\{X^n, n \ge 1\}$ is $C([0,\infty), S_0)$-exponentially tight of order a_n if and only if the following conditions hold:*

1. *the sequence $\{X^n(t), n \ge 1\}$ is S_0-exponentially tight of order a_n for all $t \ge 0$,*

2. $$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P^{1/a_n}\left(\sup_{\substack{|s-t|\le\delta \\ s,t\le T}} d(X^n(t), X^n(s)) > \eta \right) = 0, \eta > 0, T > 0.$$

For the case $S = (\mathcal{V}^+_b)^J$ and $S_0 = (\mathcal{V}^+_{b,c,1})^J$, where $J = 1, 2, \dots$, conditions 1 and 2 take the form

1'. a) $$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} P^{1/a_n}\left(\sup_{\substack{x>r, \\ y>r}} |X^n_j(x,t) - X^n_j(y,t)| > \eta \right) = 0, \ \eta > 0, t \ge 0, j = 1, ..., J,$$

b) $$\lim_{n\to\infty} P^{1/a_n}\left(\sup_{\substack{x\le r, \\ y\le r}} (|X^n_j(x,t) - X^n_j(y,t)| - |x-y|) > \eta \right) = 0, \ r > 0, \eta > 0, t \ge 0, j = 1, ..., J,$$

2'. $$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P^{1/a_n}\left(\sup_{\substack{|t-s|\le\delta, \\ s,t\le T}} \rho(X^n_j(t), X^n_j(s)) > \eta \right) = 0, \ \eta > 0, T > 0, j = 1, ..., J.$$

The proof is analogous to that in Ethier and Kurtz [7]. The second part follows by the definition of the topology on $\mathcal{V}^+_b$.

3 Fluid model. Main result.

We consider the queueing network introduced in Section 1. Let $s(l)$ denote the node that serves class l customers, where $l = 1, 2, \dots, L$, and $c(m), m = 1, \dots, M$, denote the set of customer classes that are served by node m: $c(m) = \{l \in \{1, 2, \dots, L\} : s(l) = m\}$. We characterise the state of the network at time t by the vector $N(t) = ((N_l(k,t), k = 1, 2, ...), l = 1, 2, ..., L)$, where $N_l(k,t)$ denotes the number of class l customers among the first k customers in line to $s(l)$ at t (recall that the customers at a node form a line in order of their arrivals). For $l = 1, ..., L$, we denote by $E_l(t)$ the number of exogenous class l arrivals by t and by $S_l(t)$, the number of class l customers served by node $s(l)$ during the first t units of time of serving this class customers. For $l, l' = 1, 2, ..., L$, we denote by $\Phi_{ll'}(i)$ the number of class l customers out of the first i class l customers served by $s(l)$ who become class l' customers after the service is completed. All the random objects are assumed to be defined on a probability space $(\Omega, \mathcal{F}, P)$.

We now introduce a fluid model associated with the network. To this end, we consider a sequence of networks indexed by n that differ from the above network only by initial states and have the same arrival, service and class-change processes, i.e., specified by the respective random variables $E_l(t)$, $S_l(t)$ and $\Phi_{ll'}(i)$. The state of the nth network at time t is denoted by $N^n(t) = (N^n_l(k,t), l = 1,...,L, k = 1,2,...)$. The associated normalised and time-scaled processes are defined by

$$
\begin{aligned}
E^n_l(t) &= \frac{1}{n}E_l(nt), E^n(t) = (E^n_l(t), l = 1,...,L), E^n = (E^n(t), t \geq 0),\\
S^n_l(t) &= \frac{1}{n}S_l(nt), S^n(t) = (S^n_l(t), l = 1,...,L), S^n = (S^n(t), t \geq 0),\\
\Phi^n_{ll'}(t) &= \frac{1}{n}\Phi_{ll'}(\lfloor nt \rfloor), \Phi^n(t) = (\Phi^n_{ll'}(t), l, l' = 1,...,L), \Phi^n = (\Phi^n(t), t \geq 0),\\
F^n_l(x,t) &= \frac{1}{n}N^n_l(\lfloor nx \rfloor, nt), F^n_l(t) = (F^n_l(x,t), x \geq 0),\\
F^n(t) &= (F^n_l(t), l = 1,...,L), F^n = (F^n(t), t \geq 0).
\end{aligned}
\tag{3.1}
$$

The processes E^n, S^n and Φ^n are regarded as random elements of the respective spaces $D^+([0,\infty), \mathbf{R}^L)$, $D^+([0,\infty), \mathbf{R}^L)$ and $D^+([0,\infty), \mathbf{R}^{L\times L})$. Also, we regard $F^n(t)$ as a random element of the space $(\mathcal{V}^+_b)^L$ and F^n, as a random element of the Skorohod space $D([0,\infty), (\mathcal{V}^+_b)^L)$. Note that the number of class l customers queued up at node $s(l)$ at time t is $Q^n_l(t) = N^n_l(\infty, t)$.

We next introduce processes $B^n = (B^n(t), t \geq 0)$ by $B^n(t) = (B^n_l(t), l = 1,...,L)$, $B^n_l(t) = \bar{B}^n_l(nt)/n$, where $\bar{B}^n_l(t)$ denotes the cumulative time spent by node $s(l)$ serving class l customers by time t in the nth network. Since the increments $\sum_{l\in c(m)} \overline{B}^n_l(t) - \sum_{l\in c(m)} \overline{B}^n_l(s), m = 1,\ldots,M$, for $t \geq s$ are not greater than $t - s$, the B^n are regarded as random elements of $\mathcal{V}^+_{c,1,L}$, the subspace of $C^+([0,\infty), \mathbf{R}^L)$ of functions $f = (f_1,\ldots,f_L)$ such that $\sum_{l\in c(m)} f_l \in \mathcal{V}^+_{c,1}, m = 1,\ldots,M$, endowed with relative topology.

For $t_0 \geq 0$, we define shifted processes $F^{n,t_0} = (F^{n,t_0}(t), t \geq 0)$, $B^{n,t_0} = (B^{n,t_0}(t), t \geq 0)$, $E^{n,t_0} = (E^{n,t_0}(t), t \geq 0)$, $S^{n,t_0} = (S^{n,t_0}(t), t \geq 0)$ and $\Phi^{n,t_0} = (\Phi^{n,t_0}(t), t \geq 0)$ by

$$
\begin{aligned}
F^{n,t_0}(t) = F^n(t + t_0),\ \ B^{n,t_0}(t) = B^n(t + t_0) - B^n(t_0),\ \ E^{n,t_0}(t) = E^n(t + t_0) - E^n(t_0),\\
S^{n,t_0}_l(t) = S^n_l(t + B^n_l(t_0)) - S^n_l(B^n_l(t_0)),\ l = 1,2,\ldots,L,\\
\Phi^{n,t_0}_{ll'}(t) = \Phi^n_{ll'}(t + S^n_l(B^n_l(t_0))) - \Phi^n_{ll'}(S^n_l(B^n_l(t_0))),\ l,l' = 1,2,\ldots,L.
\end{aligned}
$$

Let us assume that the following laws of large numbers hold:

$$
E^n_l(t) \xrightarrow{P} \lambda_l t,\ S^n_l(t) \xrightarrow{P} \mu_l t,\ \Phi^n_{ll'}(t) \xrightarrow{P} p_{ll'}t,\ \ l,l' = 1,2,\ldots,L, \tag{3.2}
$$

where $\xrightarrow{P}$ denotes convergence in probability. We denote

$$
\lambda = ((\lambda_l t, l = 1,...,L), t \geq 0),\ \mu = ((\mu_l t, l = 1,...,L), t \geq 0),\ p = ((p_{ll'}t, l, l' = 1,...,L), t \geq 0).
$$

Let $\mathcal{V}^+_{b,c,1,L}$ be the subspace of $(\mathcal{V}^+_{b,c,1})^L$ of functions $f = (f_1,\ldots,f_L)$ such that

$$
\sum_{l\in c(m)} f_l(x) = x \text{ for } x \leq \sum_{l\in c(m)} \|f_l\|,\ m = 1,\ldots,M. \tag{3.3}
$$

Note that $\mathcal{V}^+_{b,c,1,L}$ is closed in $(\mathcal{V}^+_b)^L$.

Let $G = (G_t(f, \mathbf{b}, \mathbf{e}, \mathbf{s}, \phi))$, where $t \geq 0$, $f \in D([0,\infty), (\mathcal{V}^+_b)^L)$, $\mathbf{b} \in \mathcal{V}^+_{c,1,L}$, $\mathbf{e} \in D^+([0,\infty), \mathbf{R}^L)$, $\mathbf{s} \in D^+([0,\infty), \mathbf{R}^L)$, and $\phi \in D^+([0,\infty), \mathbf{R}^{L\times L})$ be a function with values in a finite-dimensional vector space, which is $C([0,\infty), \mathcal{V}^+_{b,c,1,L}) \times \mathcal{V}^+_{c,1,L} \times C^+([0,\infty), \mathbf{R}^L) \times C^+([0,\infty), \mathbf{R}^L) \times$

$C^+([0,\infty),\mathbf{R}^{L\times L})$-continuous and Borel-measurable in $f, \mathbf{b}, \mathbf{e}, \mathbf{s}$, and ϕ for every t. We also assume that G has the following self-similarity property:

$$\frac{1}{k}\, G_{kt}(f,\mathbf{b},\lambda,\mu,p) = G_t(f*k, b*k, \lambda,\mu,p),\, t\ge 0, k=1,2,\dots, \tag{3.4}$$

where we denote $f*k = (k^{-1}f(kx,kt), x\ge 0, t\ge 0)$ and $b*k = (k^{-1}b(kt), t\ge 0)$.

We say that G specifies a fluid model if for all $n=1,2,\dots$, initial states $N^n(0)$ and $t_0\ge 0$,

$$G_t(F^{n,t_0}, B^{n,t_0}, E^{n,t_0}, S^{n,t_0}, \Phi^{n,t_0}) = 0,\; t\ge 0. \tag{3.5}$$

Let $\mathrm{Lip}_\beta([0,\infty),\mathcal{V}^+_{\mathrm{b,c,1,L}})$, where $\beta>0$, denote the subset of $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$ of functions $f=(f_1,\dots,f_L)$ such that

$$\sup_{x\ge 0}|f_l(x,t)-f_l(x,s)| \le \beta(t-s),\; 0\le s\le t, l=1,\dots,L. \tag{3.6}$$

For $f_0\in\mathcal{V}^+_{b,c,1,L}$, we define $\mathcal{M}(f_0)$ as the set of functions $f\in\mathrm{Lip}_\beta([0,\infty),\mathcal{V}^+_{\mathrm{b,c,1,L}})$, where $\beta=\sum_{l=1}^L(\lambda_l+2\mu_l)$, with $f(0)=f_0$, for which there exist functions $\mathbf{b}\in\mathcal{V}^+_{c,1,L}$ such that

$$G_t(f,\mathbf{b},\lambda,\mu,p)=0,\; t\ge 0.$$

We also denote by $\mathcal{M}_t(f_0)$ the set $\{f(t): f\in\mathcal{M}(f_0)\}$. We refer to elements of the union $\mathcal{M}=\bigcup\{\mathcal{M}(f_0): f_0\in\mathcal{V}^+_{b,c,1,L}\}$ as to fluid solutions (trajectories) associated with G. Note that $\mathcal{M}$ is closed in $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$.

Let, for $\alpha>0$ and $g\in(\mathcal{V}^+_b)^L$, $g\circ\alpha$ denote the function $g\circ\alpha(x)=g(\alpha x), x\ge 0$; accordingly, if Γ is a set of functions, we let $\Gamma\circ\alpha := \{g\circ\alpha:\ g\in\Gamma\}$.

In view of the continuity and self-similarity properties of the function G and the fact that the set of functions $f\in\mathrm{Lip}_\beta([0,\infty),\mathcal{V}^+_{\mathrm{b,c,1,L}})$, such that the $\|f(0)\|$ belong to a compact set, is compact in $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$ (see the proof of Lemma 5.1 for details), the collection $\mathcal{M}$ has the following properties:

(self-similarity) for $k=1,2,\ldots$ and $f_0\in\mathcal{V}^+_{b,c,1,L}$,

$$\mathcal{M}(k^{-1}f_0\circ k) = \{f*k,\, f\in\mathcal{M}(f_0)\},$$

in particular,

$$\mathcal{M}_t\left(\frac{1}{k}f_0\circ k\right) = \frac{1}{k}\mathcal{M}_{kt}(f_0)\circ k,\; t\ge 0,$$

(compactness) for every $a>0$ the set $\bigcup_{f_0\in\mathcal{V}^+_{b,c,1,L}:\|f_0\|\le a}\mathcal{M}(f_0)$ is compact in $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$.

As it follows from Lemmas 4.1 and 5.1 below, these conditions are natural to impose on a fluid model.

Next, we assume the following condition for the processes E, S and Φ.

(C) The sequences $\{E^n, n\ge 1\}$, $\{S^n, n\ge 1\}$ and $\{\Phi^n, n\ge 1\}$ are, respectively, $C^+([0,\infty),\mathbf{R}^L)$-, $C^+([0,\infty),\mathbf{R}^L)$- and $C^+([0,\infty),\mathbf{R}^{L\times L})$-exponentially tight of order n. If rate functions I_E, I_S and I_Φ are respective LD accumulation points of $\{E^n, n\ge 1\}$, $\{S^n, n\ge 1\}$ and $\{\Phi^n, n\ge 1\}$ for the normalising sequence n, then they equal 0 only at the functions λ, μ and p, respectively.

Condition (C) holds, e.g., in the standard setting when for each customer class the exogenous interarrival times and service times are i.i.d. and switching between classes is Bernoulli. This follows, e.g., from the results of Puhalskii and Whitt [15]. Note that condition (C) implies (3.2).

Our main result is the following theorem.

Theorem 1 *Let function $\hat{f}_0 \in \mathcal{V}^+_{b,c,1,L}$ be such that*

$$C := \underline{\lim}_{t\to\infty} \inf_{\hat{f}\in\mathcal{M}(\hat{f}_0)} \frac{||\hat{f}(t)||}{t} > 0.$$

Let there exist a $\mathcal{V}^+_{b,c,1,L}\times\mathcal{V}^+_{b,c,1,L}$-continuous function $R : (\mathcal{V}^+_b)^L\times\mathcal{V}^+_{b,c,1,L} \to \mathbf{R}_+$ such that $R(f,g) \geq |||g|| - ||f|||$ and $R(\alpha f, \alpha g) = \alpha R(f,g)$ for $f \in (\mathcal{V}^+_b)^L, g \in \mathcal{V}^+_{b,c,1,L}, \alpha > 0$, integer k_0 and reals $\varepsilon \in (0, C)$, $\eta \in (0, C/\varepsilon - 1)$ and $\delta > 0$, for which the following conditions hold:

1. *if function $f_0 \in \mathcal{V}^+_{b,c,1,L}$ is such that $R(f_0, \hat{f}_0) \leq \delta$, then*

$$\sup_{g\in\mathcal{M}_{2^{k_0}}(f_0)} \inf_{\tilde{g}\in\mathcal{M}_{2^{k_0}}(\hat{f}_0)} R(g\circ 2^{k_0}, \tilde{g}\circ 2^{k_0}) \leq \varepsilon 2^{k_0};$$

2. *if function $f_0 \in \mathcal{V}^+_{b,c,1,L}$ is such that, for some $k \in \{k_0, k_0+1, ...\}$,*

$$\inf_{\tilde{g}\in\mathcal{M}_{2^k}(\hat{f}_0)} R(f_0\circ 2^k, \tilde{g}\circ 2^k) \leq (1+\eta)\varepsilon 2^k,$$

   then

$$\sup_{g\in\mathcal{M}_{2^k}(f_0)} \inf_{\tilde{g}\in\mathcal{M}_{2^{k+1}}(\hat{f}_0)} R(g\circ 2^{k+1}, \tilde{g}\circ 2^{k+1}) \leq \varepsilon 2^{k+1}.$$

Let, in addition, ess sup$_{\omega\in\Omega} R(F^n(0), \hat{f}_0) \to 0$ as $n \to \infty$.

Then

$$\lim_{n\to\infty} P\left(\underline{\lim}_{k\to\infty} \frac{||F^n(2^k)||}{2^k} > C - (1+\eta)\varepsilon \right) = 1.$$

Corollary 1 *Under the hypotheses of Theorem 1, we have for all n large enough*

$$\lim_{a\to\infty} \overline{\lim}_{t\to\infty} P(||N^n(t)|| \geq a) > 0,$$

i.e., the process $N^n(t), t \geq 0$, is non-ergodic.

Proof. Since for arbitrary $\gamma > 0$

$$\overline{\lim}_{t\to\infty} P(||N^n(t)|| \geq a) \geq \overline{\lim}_{k\to\infty} P\left(\frac{||F^n(2^k)||}{2^k} \geq \gamma\right) \geq P\left(\underline{\lim}_{k\to\infty} \frac{||F^n(2^k)||}{2^k} \geq \gamma\right),$$

the corollary follows by the fact that, in view of Theorem 1,

$$P\left(\underline{\lim}_{k\to\infty} \frac{||F^n(2^k)||}{2^k} > C - (1+\eta)\varepsilon \right) > 0$$

for all n large enough.

4 Proof of Theorem 1.

We start with two simple lemmas. For an integer-valued sequence $\{m_n, n \geq 1\}$, where $m_n \geq 1$, we denote $F^{n(m_n)}(t) = m_n^{-1} F^n(m_n(1+t)) \circ m_n, t \geq 0$, and $F^{n(m_n)} = (F^{n(m_n)}(t), t \geq 0)$. The above notation is preserved.

Lemma 4.1 *Let the sequence $\{\|F^n(0)\|, n = 1, 2, ...\}$ be exponentially tight of order nm_n in $\mathbf{R}_+$. Then the sequence $\{F^{n(m_n)}, n = 1, 2, \ldots\}$ is $C([0,\infty), \mathcal{V}^+_{b,c,1,L})$-exponentially tight of order nm_n in $D([0,\infty), (\mathcal{V}^+_b)^L)$.*

Proof. We first prove that the sequence is $C([0,\infty), (\mathcal{V}^+_{b,c,1})^L)$-exponentially tight. By Lemma 2.3 and the obvious inequality $\rho(f,g) \leq \sup_{x \geq 0} |f(x) - g(x)|$, $f, g \in \mathcal{V}^+_b$, the required is implied by the following:

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} P^{1/nm_n} \left(\sup_{\substack{x>r,\\ y>r}} |F_l^{n(m_n)}(x,t) - F_l^{n(m_n)}(y,t)| > \eta \right) = 0, \ \eta > 0, t \geq 0, l = 1, ..., L, \quad (4.1)$$

$$\lim_{n\to\infty} P^{1/nm_n} \left(\sup_{\substack{x\leq r,\\ y\leq r}} (|F_l^{n(m_n)}(x,t) - F_l^{n(m_n)}(y,t)| - |x-y|) > \eta \right) = 0, \ r > 0, \eta > 0, t \geq 0, l = 1, ..., L, \quad (4.2)$$

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P^{1/nm_n} \left(\sup_{\substack{|t-s|\leq\delta,\\ s,t\leq T}} \sup_{x\geq 0} |F_l^{n(m_n)}(x,t) - F_l^{n(m_n)}(x,s)| > \eta \right) = 0, \ \eta > 0, T > 0, l = 1, ..., L \quad (4.3)$$

For (4.1), note that since $N_l^n(k,t) = \|N_l^n(t)\|$ for $k \geq \sum_{l'=1}^L Q_l^n(0) + \sum_{l'=1}^L E_l(t)$, we have that

$$F^{n(m_n)}(x,t) = F^{n(m_n)}(y,t) \ \text{ if } x, y \geq m_n^{-1} \|F^n(0)\| + \sum_{l=1}^L E_l^{nm_n}(1+t).$$

The required now follows by exponential tightness of $\{\|F^n(0)\|, n \geq 1\}$ of order nm_n, the inequality $m_n \geq 1$, and the fact that, by exponential tightness of $\{E^n, n \geq 1\}$ of order n and Lemma 2.3,

$$\lim_{A\to\infty} \overline{\lim_{n\to\infty}} P^{1/nm_n}(E_l^{nm_n}(1+t) > A) = 0, \ l = 1, 2, \ldots, L.$$

Limit (4.2) holds, since

$$|F_l^{n(m_n)}(x,t) - F_l^{n(m_n)}(y,t)| = \frac{1}{nm_n} |N_l^n(\lfloor nm_n x \rfloor, nm_n(1+t)) - N_l^n(\lfloor nm_n y \rfloor, nm_n(1+t))| \leq \frac{|\lfloor nm_n x \rfloor - \lfloor nm_n y \rfloor|}{nm_n}.$$

To prove (4.3), let us denote, given an initial state $N_l^n(0) = nF_l^n(0), l = 1, \ldots, L$, and $k = 1, 2, ...$:

$\bar{A}_l^n(t), l = 1, ..., L$, the number of class l arrivals at node $s(l)$ by t,

$\bar{D}_{l,k}^n(t), l = 1, ..., L$, the number of class l customers, served by node $s(l)$ by time t, who occupy positions from 1 to k in the queue,

$\bar{D}_{(l),k}^n(t), l = 1, ..., L$, the total number of customers, except class l customers, served by node $s(l)$ by time t, who occupy positions from 1 to k in the queue,

$\chi^n_{l,k}(t)$, the indicator of the event that the $(k+1)$st position in the queue to $s(l)$ is occupied at time t by a class l customer.

We obviously have

$$N_l^n(k,t) = N_l^n(k,0) + \int_0^t 1\left(\sum_{l'\in c(s(l))} Q^n_{l'}(s-) \le k-1\right) d\bar{A}^n_l(s) + \int_0^t \chi^n_{l,k}(s-)d\bar{D}^n_{(l),k}(s)$$
$$- \int_0^t (1-\chi^n_{l,k}(s-))d\bar{D}^n_{l,k}(s).$$

Thus, for $0 \le s < t$,

$$|N_l^n(k,t) - N_l^n(k,s)| \le (\bar{A}^n_l(t) - \bar{A}^n_l(s)) + (\bar{D}^n_{(l),k}(t) - \bar{D}^n_{(l),k}(s)) + (\bar{D}^n_{l,k}(t) - \bar{D}^n_{l,k}(s)).$$

Let $\bar{D}^n_l(t)$ denote the number of class l customers served by time t. Then

$$\bar{A}^n_l(t) - \bar{A}^n_l(s) \le (E_l(t) - E_l(s)) + \sum_{l'=1}^{L}(\bar{D}^n_{l'}(t) - \bar{D}^n_{l'}(s)),$$

$$\bar{D}^n_{l,k}(t) - \bar{D}^n_{l,k}(s) \le \bar{D}^n_l(t) - \bar{D}^n_l(s), \;\; \bar{D}^n_{(l),k}(t) - \bar{D}^n_{(l),k}(s) \le \sum_{\substack{l'=1\\ l'\ne l}}^{L}(\bar{D}^n_{l'}(t) - \bar{D}^n_{l'}(s)).$$

Hence,

$$|N_l^n(k,t) - N_l^n(k,s)| \le (E_l(t) - E_l(s)) + 2\sum_{l'=1}^{L}(\bar{D}^n_{l'}(t) - \bar{D}^n_{l'}(s)),$$

so that, in view of (3.1) and the equality $\bar{D}^n_{l'}(t) = S_{l'}(\bar{B}^n_{l'}(t))$,

$$|F^n_l(x,t) - F^n_l(x,s)| \le (E^n_l(t) - E^n_l(s)) + 2\sum_{l'=1}^{L}(S^n_{l'}(B^n_{l'}(t)) - S^n_{l'}(B^n_{l'}(s))).$$

Therefore, since $(t - B^n_{l'}(t))$ is nondecreasing,

$$\sup_{\substack{|s-t|\le\delta\\ s,t\le T}} \sup_{x\ge 0} |F^n_l(x,t) - F^n_l(x,s)| \le \sup_{\substack{|s-t|\le\delta\\ s,t\le T}} |E^n_l(t) - E^n_l(s)| + 2\sum_{l'=1}^{L} \sup_{\substack{|s-t|\le\delta\\ s,t\le T}} |S^n_{l'}(t) - S^n_{l'}(s)|. \quad (4.4)$$

Condition (C) and Lemma 2.3 easily imply that, for $\eta > 0$,

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P^{1/n}\left(\sup_{\substack{|s-t|\le\delta,\\ s,t\le T}} |E^n_l(t) - E^n_l(s)| > \eta\right) = 0,$$

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P^{1/n}\left(\sup_{\substack{|s-t|\le\delta,\\ s,t\le T}} |S^n_l(t) - S^n_l(s)| > \eta\right) = 0, l = 1, ..., L.$$

Thus, (4.4) implies (4.3).

Therefore, the sequence $\{F^{n(m_n)}, n=1,2,\ldots\}$ is $C([0,\infty),(\mathcal{V}^+_{b,c,1})^L)$-exponentially tight. The $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$-exponential tightness follows now by the fact that

$$\sum_{l\in c(m)} F_l^{n(m_n)}(x,t)=\frac{\lfloor nm_nx\rfloor}{nm_n} \text{ if } \frac{\lfloor nm_nx\rfloor}{nm_n}\le \sum_{l\in c(m)}\|F_l^{n(m_n)}(t)\|,\, t\ge 0, m=1,\ldots,M$$

(cf., the proof of Lemma 4.2 below). The lemma is proved.

The next lemma shows that "zeros" of LD accumulation points of $\{F^{n(m_n)}, n\ge 1\}$ are fluid solutions associated with G. For $f_0\in\mathcal{V}^+_{b,c,1,L}$, we denote by $\mathcal{A}(f_0)$ the set of all functions $f=(f(t),t\ge 0)$ starting at f_0, for which there exist a sequence m_n and a $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$–exponentially tight of order nm_n sequence $\{F^{n(m_n)}, n\ge 1\}$ such that some LD accumulation point of $\{F^{n(m_n)}, n\ge 1\}$ for the normalising sequence nm_n equals zero at f. Accordingly, $\mathcal{A}_t(f_0)$ denotes the set of the values of functions from $\mathcal{A}(f_0)$ at t: $\mathcal{A}_t(f_0)=\{f(t):\ f\in\mathcal{A}(f_0)\}$. Note that the set $\mathcal{A}(f_0)$ is non-empty by Lemma 4.1. Let also $\mathcal{A}=\bigcup\{\mathcal{A}(f_0):\ f_0\in\mathcal{V}^+_{b,c,1,L}\}$.

Lemma 4.2 *$\mathcal{A}(f_0)\subset\mathcal{M}(f_0)$ for $f_0\in\mathcal{V}^+_{b,c,1,L}$.*

Proof. Let $\tilde f\in\mathcal{A}(f_0)$ and $\tilde I^F(f)$ be an LD accumulation point for a normalising sequence nm_n of a $C([0,\infty),\mathcal{V}^+_{b,c,1,L})$-exponentially tight of order nm_n sequence $\{F^{n(m_n)}, n\ge 1\}$ such that $\tilde I^F(\tilde f)=0$. Let us denote $\tilde B^n=B^{nm_n,1}, \tilde E^n=E^{nm_n,1}, \tilde S^n=S^{nm_n,1}, \tilde\Phi^n=\Phi^{nm_n,1}$. Since the trajectories of B^n are from $\mathcal{V}^+_{c,1,L}$, Lemma 2.3 implies that the sequence $\{\tilde B^n, n\ge 1\}$ is exponentially tight of order nm_n. Similarly, it is not difficult to see by condition (C) and the inequality $B^n(t)\le t$ that the sequences $\{\tilde E^n, n\ge 1\}$, $\{\tilde S^n, n\ge 1\}$ and $\{\tilde\Phi^n, n\ge 1\}$ are, respectively, $C^+([0,\infty),\mathbf{R}^L)$-, $C^+([0,\infty),\mathbf{R}^L)$- and $C^+([0,\infty),\mathbf{R}^{L\times L})$-exponentially tight of order nm_n. For instance, condition 2 of Lemma 2.3 holds for $\{\tilde\Phi^n, n\ge 1\}$, since for $\delta>0,\eta>0,T>0$, and $T_1>0$,

$$P\left(\sup_{\substack{|s-t|\le\delta\\ s,t\le T}}|\tilde\Phi^n_{ll'}(t)-\tilde\Phi^n_{ll'}(s)|>\eta\right)$$

$$=P\left(\sup_{\substack{|s-t|\le\delta\\ s,t\le T}}\left|\Phi^{nm_n}_{ll'}\left(t+S^{nm_n}_l\left(B^{nm_n}_l(1)\right)\right)-\Phi^{nm_n}_{ll'}\left(s+S^{nm_n}_l\left(B^{nm_n}_l(1)\right)\right)\right|>\eta\right)$$

$$\le P\left(\sup_{\substack{|s-t|\le\delta\\ s,t\le T_1}}|\Phi^{nm_n}_{ll'}(t)-\Phi^{nm_n}_{ll'}(s)|>\eta\right)+P(S^{nm_n}_l(1)+T>T_1),$$

and the sequences $\{\Phi^{nm_n}, n\ge 1\}$ and $\{S^{nm_n}, n\ge 1\}$ are, respectively, $C^+([0,\infty),\mathbf{R}^{L\times L})$- and $C^+([0,\infty),\mathbf{R}^L)$-exponentially tight of order nm_n.

Lemma 2.3 then implies, that the sequence $\{(F^{n(m_n)},\tilde B^n,\tilde E^n,\tilde S^n,\tilde\Phi^n), n\ge 1\}$ is $C([0,\infty),\mathcal{V}^+_{b,c,1,L})\times\mathcal{V}^+_{c,1,L}\times C^+([0,\infty),\mathbf{R}^L)\times C^+([0,\infty),\mathbf{R}^L)\times C^+([0,\infty),\mathbf{R}^{L\times L})$-exponentially tight of order nm_n. Let a subsequence n' be such that the sequences $\{F^{n'(m_{n'})}\}$ and $\{(F^{n'(m_{n'})},\tilde B^{n'},\tilde E^{n'},\tilde S^{n'},\tilde\Phi^{n'})\}$ obey LDPs for the normalising sequence $n'm_{n'}$ with respective rate functions $\tilde I^F$ and $\tilde I$. By the contraction principle, $\tilde I^F(\tilde f)=\inf_{\mathbf{b},\mathbf{e},\mathbf{s},\phi}\tilde I(\tilde f,\mathbf{b},\mathbf{e},\mathbf{s},\phi)$. Since $\tilde I$ is a rate function, the infimum is attained at some functions $\tilde{\mathbf{b}}\in\mathcal{V}^+_{c,1,L}$, $\tilde{\mathbf{e}}\in C^+([0,\infty),\mathbf{R}^L)$, $\tilde{\mathbf{s}}\in C^+([0,\infty),\mathbf{R}^L)$ and $\tilde\phi\in C^+([0,\infty),\mathbf{R}^{L\times L})$; hence, $\tilde I(\tilde f,\tilde{\mathbf{b}},\tilde{\mathbf{e}},\tilde{\mathbf{s}},\tilde\phi)=0$. On the other hand, applying the contraction principle once again, we have that $\{\tilde{E}^{n'}\}$, $\{\tilde{S}^{n'}\}$ and $\{\tilde{\Phi}^{n'}\}$ obey LDPs in the respective spaces $D^+([0,\infty),\mathbf{R}^L)$, $D^+([0,\infty),\mathbf{R}^L)$ and $D^+([0,\infty),\mathbf{R}^{L\times L})$ for the normalising sequence $n'm_{n'}$ with the respective rate functions $I'_E(\mathbf{e}) = \inf_{f,\mathbf{b},\mathbf{s},\phi} \tilde{I}(f,\mathbf{b},\mathbf{e},\mathbf{s},\phi)$, $I'_S(\mathbf{s}) = \inf_{f,\mathbf{b},\mathbf{e},\phi} \tilde{I}(f,\mathbf{b},\mathbf{e},\mathbf{s},\phi)$ and $I'_\Phi(\phi) = \inf_{f,\mathbf{b},\mathbf{e},\mathbf{s}} \tilde{I}(f,\mathbf{b},\mathbf{e},\mathbf{s},\phi)$. Therefore, $I'_E(\tilde{\mathbf{e}}) = 0$, $I'_S(\tilde{\mathbf{s}}) = 0$ and $I'_\Phi(\tilde{\phi}) = 0$, and by condition (C) $\tilde{\mathbf{e}} = \lambda$, $\tilde{\mathbf{s}} = \mu$, and $\tilde{\phi} = p$. Let us prove, say, the second equality. Since by condition (C) the sequence $\{S^n, n \geq 1\}$ is $C^+\left([0,\infty),R^L\right)$-exponentially tight of order n, we can assume, considering a subsequence of n' if necessary, that the sequence $\{S^{n'm_{n'}}\}$ obeys an LDP for the normalising sequence $n'm_{n'}$ with some rate function I''_S. Then by Lemma 2.2 , for $T > 0$,

$$1 = \lim_{\varepsilon\to 0} \overline{\lim_{n'\to\infty}} P^{1/n'm_{n'}}\left(\sup_{t\leq T} |\tilde{S}^{n'}(t) - \tilde{\mathbf{s}}(t)| \leq \varepsilon\right)$$

$$= \lim_{\varepsilon\to 0} \overline{\lim_{n'\to\infty}} P^{1/n'm_{n'}}\left(\sup_{t\leq T}\left|\left(S^{n'm_{n'}}\left(t + B^{n'm_{n'}}(1)\right) - S^{n'm_{n'}}\left(B^{n'm_{n'}}(1)\right)\right) - \tilde{\mathbf{s}}(t)\right| \leq \varepsilon\right)$$

$$\leq \lim_{\varepsilon\to 0} \overline{\lim_{n'\to\infty}} P^{1/n'm_{n'}}\left(\sup_{t\leq T}\inf_{0\leq x\leq 1}\left|\left(S^{n'm_{n'}}(t+x) - S^{n'm_{n'}}(x)\right) - \tilde{\mathbf{s}}(t)\right| \leq \varepsilon\right) = \exp\left(-\inf_{\mathbf{s}\in\Theta} I'_S(\mathbf{s})\right),$$

where $\Theta = \{\mathbf{s} \in C([0,\infty),\mathbf{R}^L) : \sup_{t\leq T}\inf_{0\leq x\leq 1}|\mathbf{s}(t+x) - \mathbf{s}(x) - \tilde{\mathbf{s}}(t)| = 0\}$. Condition (C) then implies that $\mu \in \Theta$ so that for every t there exists an x such that $\tilde{\mathbf{s}}(t) = \mu(t+x) - \mu(x) = \mu(t)$.

Since by (3.5) $G_t(F^{n'(m_{n'})}, \tilde{B}^{n'}, \tilde{E}^{n'}, \tilde{S}^{n'}, \tilde{\Phi}^{n'}) = 0$, $t \geq 0$, and the function $G_t(\cdot)$, for given t, is $C([0,\infty),\mathcal{V}^+_{b,c,1,L}) \times \mathcal{V}^+_{c,1,L} \times C^+([0,\infty),\mathbf{R}^L) \times C^+([0,\infty),\mathbf{R}^L) \times C^+([0,\infty),\mathbf{R}^{L\times L})$-continuous, by the extended contraction principle $G_t(f,\mathbf{b},\mathbf{e},\mathbf{s},\phi) = 0$ when $\tilde{I}(f,\mathbf{b},\mathbf{e},\mathbf{s},\phi) < \infty$. Since $\tilde{I}(\tilde{f},\tilde{\mathbf{b}},\lambda,\mu,p) = 0 < \infty$, it follows that $G_t(\tilde{f},\tilde{\mathbf{b}},\lambda,\mu,p) = 0$.

By a similar reasoning the relation (4.4) implies that, for $\delta > 0$ and $T > 0$,

$$\sup_{\substack{|s-t|\leq\delta \\ s,t\leq T}} \sup_{x\geq 0} |\tilde{f}_l(x,t) - \tilde{f}_l(x,s)| \leq \lambda_l\delta + 2\sum_{l'=1}^{L}\mu_{l'}\delta$$

so that $\tilde{f}$ satisfies (3.6) with $\beta = \sum_{l=1}^{L}(\lambda_l + 2\mu_l)$. Thus, $\tilde{f} \in \mathcal{M}(f_0)$. The lemma is proved.

Proof of Theorem 1. Let us denote

$$R(f,\Gamma) = \inf_{g\in\Gamma} R(f,g), f \in (\mathcal{V}^+_b)^L, \Gamma \subset \mathcal{V}^+_{b,c,1,L}.$$

Since

$$\underline{\lim}_{t\to\infty} \inf_{\hat{f}\in\mathcal{M}(\hat{f}_0)} \frac{||\hat{f}(t)||}{t} = C,$$

the inequality

$$||f|| - ||F^n(t)|| \leq R(F^n(t)\circ t, f\circ t), f \in \mathcal{V}^+_{b,c,1,L},$$

allows us to write

$$P\left(\underline{\lim}_{k\to\infty} \frac{||F^n(2^k)||}{2^k} \leq C - (1+\eta)\varepsilon\right)$$

$$\leq P\left(\underline{\lim}_{k\to\infty} \frac{||F^n(2^k)||}{2^k} \leq \underline{\lim}_{k\to\infty} \inf_{\hat{f}\in\mathcal{M}(\hat{f}_0)} \frac{||\hat{f}(2^k)||}{2^k} - (1+\eta)\varepsilon\right)$$

$$\le P\left(\overline{\lim_{k\to\infty}} \inf_{\hat f \in \mathcal{M}(\hat f_0)} \frac{||\hat f(2^k)|| - ||F^n(2^k)||}{2^k} \ge (1+\eta)\varepsilon \right)$$

$$\le P\left(\overline{\lim_{k\to\infty}} \inf_{\hat f \in \mathcal{M}(\hat f_0)} \frac{R(F^n(2^k)\circ 2^k, \hat f(2^k)\circ 2^k)}{2^k} \ge (1+\eta)\varepsilon \right)$$

$$\le P\left(\sup_{k=k_0,k_0+1,\dots} \inf_{\hat f \in \mathcal{M}(\hat f_0)} \frac{R(F^n(2^k)\circ 2^k, \hat f(2^k)\circ 2^k)}{2^k} \ge (1+\eta)\varepsilon \right).$$

Thus, it suffices to show that

$$\lim_{n\to\infty} P\left(\sup_{k=k_0,k_0+1,\dots} \frac{R\left(F^n(2^k)\circ 2^k, \mathcal{M}_{2^k}(\hat f_0)\circ 2^k\right)}{2^k} \ge (1+\eta)\varepsilon \right) = 0. \tag{4.5}$$

We have

$$P\left(\sup_{k=k_0,k_0+1,\dots} \frac{R\left(F^n(2^k)\circ 2^k, \mathcal{M}_{2^k}(\hat f_0)\circ 2^k\right)}{2^k} \ge (1+\eta)\varepsilon \right)$$
$$\le P\left(R\left(\frac{F^n(2^{k_0})\circ 2^{k_0}}{2^{k_0}}, \frac{\mathcal{M}_{2^{k_0}}(\hat f_0)\circ 2^{k_0}}{2^{k_0}} \right) \ge (1+\frac{\eta}{2})\varepsilon \right)$$
$$+ \sum_{k=k_0}^{\infty} P\left(R\left(\frac{F^n(2^k)\circ 2^k}{2^k}, \frac{\mathcal{M}_{2^k}(\hat f_0)\circ 2^k}{2^k} \right) < (1+\frac{\eta}{2})\varepsilon, \right.$$
$$\left. R\left(\frac{F^n(2^{k+1})\circ 2^{k+1}}{2^{k+1}}, \frac{\mathcal{M}_{2^{k+1}}(\hat f_0)\circ 2^{k+1}}{2^{k+1}} \right) \ge (1+\frac{\eta}{2})\varepsilon \right). \tag{4.6}$$

Let us denote

$$q = \overline{\lim_{n\to\infty}} \frac{1}{n} \sup_{k\ge k_0} \frac{1}{2^k} \log P\left(R\left(\frac{F^n(2^k)\circ 2^k}{2^k}, \frac{\mathcal{M}_{2^k}(\hat f_0)\circ 2^k}{2^k} \right) < (1+\frac{\eta}{2})\varepsilon, \right.$$
$$\left. R\left(\frac{F^n(2^{k+1})\circ 2^{k+1}}{2^{k+1}}, \frac{\mathcal{M}_{2^{k+1}}(\hat f_0)\circ 2^{k+1}}{2^{k+1}} \right) \ge (1+\frac{\eta}{2})\varepsilon \right). \tag{4.7}$$

We prove that $q < 0$. Since $R(F^n(0), \hat f_0) \to 0$ uniformly in $\omega \in \Omega$ and $\|F^n(0)\| \le \|\hat f_0\| + R(F^n(0), \hat f_0)$, by Lemma 4.1 the sequence $\{(2^{-k_n}F^n(2^{k_n}(1+t))\circ 2^{k_n}, t \ge 0), n \ge 1\}$ is $C([0,\infty), \mathcal{V}^+_{b,c,1,L})$–exponentially tight of order $n2^{k_n}$ for arbitrary $k_n \ge 0$. Hence, there exists a subsequence n' such that the sequence $\{(2^{-k_{n'}}F^{n'}(2^{k_{n'}}(1+t))\circ 2^{k_{n'}}, t \ge 0)\}$ obeys an LDP in $D([0,\infty), (\mathcal{V}^+_b)^L)$ for the normalising sequence $n'2^{k_{n'}}$ with rate function $\tilde I$ such that $\tilde I(f) = \infty$ for $f \notin C([0,\infty), \mathcal{V}^+_{b,c,1,L})$ and

$$q = \lim_{n'\to\infty} \frac{1}{n'2^{k_{n'}}} \log P\left(R\left(\frac{F^{n'}(2^{k_{n'}})\circ 2^{k_{n'}}}{2^{k_{n'}}}, \frac{\mathcal{M}_{2^{k_{n'}}}(\hat f_0)\circ 2^{k_{n'}}}{2^{k_{n'}}} \right) < (1+\frac{\eta}{2})\varepsilon, \right.$$
$$\left. R\left(\frac{F^{n'}(2^{k_{n'}+1})\circ 2^{k_{n'}+1}}{2^{k_{n'}+1}}, \frac{\mathcal{M}_{2^{k_{n'}+1}}(\hat f_0)\circ 2^{k_{n'}+1}}{2^{k_{n'}+1}} \right) \ge (1+\frac{\eta}{2})\varepsilon \right).$$

Therefore, denoting

$$A_k = \left\{ f \in D\left([0,\infty), \left(\mathcal{V}^+_b\right)^L\right) : \; R\left(f(0), \frac{\mathcal{M}_{2^k}(\hat f_0)\circ 2^k}{2^k} \right) < (1+\frac{\eta}{2})\varepsilon, \right.$$

$$R\left(f(1)\circ 2, \frac{\mathcal{M}_{2^{k+1}}(\hat{f}_0)\circ 2^{k+1}}{2^k}\right) \geq 2(1+\frac{\eta}{2})\varepsilon\right\},$$

we have by Lemma 2.1

$$q \leq -\inf_{f\in A} \tilde{I}(f),$$

where

$$A = \left(\bigcap_{l=1}^{\infty} \text{cl}\left(\bigcup_{n'\geq l} A_{k_{n'}}\right)\right) \bigcap C\left([0,\infty), \mathcal{V}^+_{b,c,1,L}\right).$$

We show that

$$f \notin \mathcal{M}(f(0)) \quad \text{for all } f \in A. \tag{4.8}$$

Let $f \in A$. Then there exists a sequence $f_{n''}$ of elements of $D\left([0,\infty), \left(\mathcal{V}^+_b\right)^L\right)$, which converges to f (in particular, $f_{n''}(0) \to f(0)$ and $f_{n''}(1) \to f(1)$) and such that $R(f_{n''}(0), 2^{-k_{n''}}\mathcal{M}_{2^{k_{n''}}}(\hat{f}_0)\circ 2^{k_{n''}}) < (1+\eta/2)\varepsilon, R(f_{n''}(1)\circ 2, 2^{-k_{n''}}\mathcal{M}_{2^{k_{n''}+1}}(\hat{f}_0)\circ 2^{k_{n''}+1}) \geq 2(1+\eta/2)\varepsilon$. In addition, since by the self-similiraty property $2^{-k}\mathcal{M}_{2^k}(\hat{f}_0)\circ 2^k = \mathcal{M}_1(2^{-k}\hat{f}_0\circ 2^k)$ and $\sup_{k\geq 1} 2^{-k}\|\hat{f}_0\circ 2^k\| < \infty$, we have by the compactness property that the set $\bigcup_{k=1}^{\infty} 2^{-k}\mathcal{M}_{2^k}(\hat{f}_0)\circ 2^k$ is relatively compact in $\mathcal{V}^+_{b,c,1,L}$. It is not difficult to deduce now, in view of homogeneity and $\mathcal{V}^+_{b,c,1,L}\times\mathcal{V}^+_{b,c,1,L}$–continuity of R, that for some $\tilde{n}$

$$R\left(f(0), 2^{-k_{\tilde{n}}}\mathcal{M}_{2^{k_{\tilde{n}}}}(\hat{f}_0)\circ 2^{k_{\tilde{n}}}\right) \leq (1+\eta)\varepsilon,$$
$$R\left(f(1)\circ 2, 2^{-k_{\tilde{n}}}\mathcal{M}_{2^{k_{\tilde{n}}+1}}(\hat{f}_0)\circ 2^{k_{\tilde{n}}+1}\right) > 2\varepsilon.$$

By condition 2 of the theorem and the self-similarity property the latter implies that

$$2^{k_{\tilde{n}}}f(1)\circ 2^{-k_{\tilde{n}}} \notin \mathcal{M}_{2^{k_{\tilde{n}}}}\left(2^{k_{\tilde{n}}}f(0)\circ 2^{-k_{\tilde{n}}}\right) = 2^{k_{\tilde{n}}}\mathcal{M}_1(f(0))\circ 2^{-k_{\tilde{n}}},$$

i.e., $f(1) \notin \mathcal{M}_1(f(0))$, as required.

Relation (4.8) implies by Lemma 4.2 that $A\bigcap\mathcal{A} = \emptyset$, so, by the definition of $\mathcal{A}$, $\tilde{I}(f) > 0$ for all $f\in A$; since the set A is closed and $\tilde{I}$ is a rate function, $\inf_{f\in A}\tilde{I}(f)$ is attained and, hence, is positive. Thus, $q < 0$.

By (4.6) and (4.7), for all n large enough,

$$P\left(\sup_{k\geq k_0}\frac{R(F^n(2^k)\circ 2^k, \mathcal{M}_{2^k}(\hat{f}_0)\circ 2^k)}{2^k} \geq (1+\eta)\varepsilon\right) \leq P\left(R\left(\frac{F^n(2^{k_0})\circ 2^{k_0}}{2^{k_0}}, \frac{\mathcal{M}_{2^{k_0}}(\hat{f}_0)\circ 2^{k_0}}{2^{k_0}}\right) \geq (1+\eta)\varepsilon\right) + \sum_{k=k_0}^{\infty}\exp(qn2^{k-1}).$$

Since $q < 0$, the second term on the right goes to 0 as $n\to\infty$. The first term goes to 0 by condition 1 of the theorem and uniform convergence of $R(F^n(0), \hat{f}_0)$ to 0. Limit (4.5) is proved. The theorem is proved.

5 Discussion. Examples.

In this section, we discuss the hypotheses of Theorem 1, state its versions, and show how the theorem applies to earlier examples of non-ergodic networks. We also discuss some properties of fluid models.

We comment first on conditions 1 and 2 of the theorem. Condition 1 holds if the fluid trajectories on bounded time intervals depend continuously on the initial state. Condition 2 means that the cone of trajectories that are close to the trajectories starting at $\hat{f}_0$ expands not faster than linearly: if at time 2^k a fluid trajectory is in a neighbourhood of size $2^k(1+\eta)\varepsilon$ about the trajectories starting at $\hat{f}_0$, then at time 2^{k+1} it will be in a neighbourhood of size $2^{k+1}\varepsilon$. The role of the metric here is played by the function R, the typical examples are $R(f,g) = |\|g\| - \|f\||$ and $R(f,g) = \sup_{x\geq 0} |g(x) - f(x)|$.

Theorem 1 gives conditions under which the process $N(t)$ is non-ergodic. If we somewhat strengthen the hypotheses by requiring that, beginning from some k_0, the trajectories that are close at time 2^k to the trajectories starting at $\hat{f}_0$ remain close to them on the entire interval $[2^k, 2^{k+1}]$, rather than only at time 2^{k+1}, we can prove that the trajectories of $N(t)$ do not return to zero. More specifically, the following statements hold, the proofs being analogous to those of Theorem 1 and Corollary 1.

Theorem 2 *Let, in the hypotheses of Theorem 1, condition 2 be replaced by the following:*

2'. if function $f_0 \in \mathcal{V}^+_{b,c,1,L}$ is such that

$$\inf_{\tilde{g}\in\mathcal{M}_{2^k}(\hat{f}_0)} R(f_0\circ 2^k, \tilde{g}\circ 2^k) \leq (1+\eta)\varepsilon 2^k$$

for some $k \in \{k_0, k_0+1, ...\}$, then

$$\sup_{g\in\mathcal{M}(f_0)} \sup_{0\leq t\leq 1} \inf_{\tilde{g}\in\mathcal{M}_{2^k}(\hat{f}_0)} (1+t)^{-1} R\left(g(2^k t)\circ(2^k(1+t)), \tilde{g}(2^k(1+t))\circ(2^k(1+t))\right) \leq \varepsilon 2^k.$$

Then

$$\lim_{n\to\infty} P\left(\underline{\lim}_{t\to\infty} \frac{\|F^n(t)\|}{t} > C - (1+\eta)\varepsilon \right) = 1.$$

Corollary 2 *Under the hypotheses of Theorem 2 for all n large enough*

$$P(\underline{\lim}_{t\to\infty} \|N^n(t)\| = \infty) > 0,$$

i.e., the process $N^n(t), t \geq 0$, is non-recurrent.

Theorems 1 and 2 impose restrictions on how close fluid trajectories are at times $2^k, k = k_0, k_0+1, \ldots$. However, the assertions carry over to the case when these times are replaced by an increasing sequence $t_k, k = k_0, k_0+1, \ldots$ such that $\overline{\lim}_{k\to\infty} t_{k+1}/t_k < \infty$ and $\underline{\lim}_{k\to\infty} t_{k+1}/t_k > 1$. The proofs differ only in notation.

The function G depends on the structure of the network and service disciplines at the nodes, and in most cases is not specified uniquely. Let us consider a typical example. Define, in addition to the above characteristics, the following (we suppress subscript n below to emphasise that the notation refers to the generic network):

$A_l(t)$, $l = 1, \ldots, L$, the number of class l arrivals at node $s(l)$ by time t,

$D_l(t)$, $l = 1, \ldots, L$, the number of class l customers served by node $s(l)$ by time t,

$I_m(t)$, $m = 1, \ldots, M$, the cumulative time that node m has been idle by time t.

We then have that for every work-conserving discipline

$$A_l(t) = E_l(t) + \sum_{l'=1}^{L} \Phi_{l'l}(D_{l'}(t)),\ D_l(t) = S_l(\overline{B}_l(t)),$$
$$Q_l(t) = Q_l(0) + A_l(t) - D_l(t),\ Q_l(t) \geq 0,\ Q_l(t) = N_l(\infty, t),$$
$$I_m(t) = t - \sum_{l \in c(m)} \overline{B}_l(t),$$
$$\overline{B}_l(t) \text{ is nondecreasing, } \overline{B}_l(0) = 0,$$
$$I_m(t) \text{ is nondecreasing, } I_m(0) = 0,$$
$$\int_0^t \left(\sum_{l \in c(m)} Q_l(s) \right) dI_m(s) = 0.$$

Thus, we may assume that the function G is specified by the following equations

$$a_l(t) = \mathbf{e}_l(t) + \sum_{l'=1}^{L} \phi_{l'l}(d_{l'}(t)),\ d_l(t) = \mathbf{s}_l(b_l(t)),\ q_l(t) = f_l(\infty, t),$$

$$q_l(t) = q_l(0) + a_l(t) - d_l(t),\ i_m(t) = t - \sum_{l \in c(m)} b_l(t),\ \int_0^t \left(\sum_{l \in c(m)} q_l(s) \right) di_m(s) = 0.$$

We obtain the fluid model equations if we let $\mathbf{e}_l(t) = \lambda_l t, \mathbf{s}_l(t) = \mu_l t, \phi_{ll'}(t) = p_{ll'} t$. All the requirements on the function G are easily met, in particular, the continuity condition with respect to $(f, \mathbf{b}, \mathbf{e}, \mathbf{s}, \phi)$ follows since $f_l(\infty), l = 1, \ldots, L$, is a continuous function of $f = (f_1, \ldots, f_L) \in (\mathcal{V}_b^+)^L$ by the definition of topology on $(\mathcal{V}_b^+)^L$.

In certain cases we might need to modify function G and condition (C) by incorporating more variables. Say, assume we want to include unfinished work into the fluid model. Denoting by $W_l(t)$, $l = 1, \ldots, L$, the cumulative unfinished work for serving class l customers at node $s(l)$ at time t, we have the equation

$$W_l(t) = U_l(Q_l(0) + A_l(t)) - \overline{B}_l(t),\ W_l(t) \geq 0,$$

where $U_l(i)$, $l = 1, \ldots, L$, denotes the time that node $s(l)$ needs for serving first i class l customers. We also have the equations

$$S_l(U_l(i)) = i,\ Q_l(t) + D_l(t) = S_l(\overline{B}_l(t) + W_l(t)).$$

If we assume the FIFO service discipline, it is characterised by the equation

$$\overline{B}_l \left(t + \sum_{l' \in c(s(l))} W_{l'}(t) \right) = \overline{B}_l(t) + W_l(t).$$

The corresponding system of "fluid" equations is

$$\mathbf{s}_l(u_l(t)) = t,\ q_l(t) + d_l(t) = \mathbf{s}_l(b_l(t) + w_l(t)),$$
$$w_l(t) = u_l(q_l(0) + a_l(t)) - b_l(t),\ b_l(t + \sum_{l' \in c(s(l))} w_{l'}(t)) = b_l(t) + w_l(t).$$

The associated function G, however, does not generally have the continuity property with respect to $(f, \mathbf{b}, \mathbf{e}, \mathbf{s}, \phi)$. We could maintain this property by adding to the processes E^n, S^n and Φ^n the process

U^n defined by $U_l^n(t) = U_l(nt)/n, U^n(t) = (U_l^n(t), l = 1, ..., L), U^n = (U^n(t), t \geq 0)$. We would then have to require existence of function $G_t(f, \mathbf{b}, \mathbf{e}, \mathbf{s}, \phi, \mathbf{u})$, where $\mathbf{u} \in D([0, \infty), \mathbf{R}^L)$, which is $C([0, \infty), (\mathcal{V}_{b,c,1}^+)^L) \times \mathcal{V}_{c,1,L}^+ \times C^+([0, \infty), \mathbf{R}^L) \times C^+([0, \infty), \mathbf{R}^L) \times C^+([0, \infty), \mathbf{R}^{L \times L}) \times C^+([0, \infty), \mathbf{R}^L)$-continuous and Borel-measurable in $f, \mathbf{b}, \mathbf{e}, \mathbf{s}, \phi$, and $\mathbf{u}$ for every t and such that for all $n = 1, 2, \ldots$, initial states $N^n(0)$ and $t_0 \geq 0$,

$$G_t(F^{n,t_0}, B^{n,t_0}, E^{n,t_0}, S^{n,t_0}, \Phi^{n,t_0}, U^{n,t_0}) = 0,\ t \geq 0,$$

where

$$U_l^{n,t_0}(t) = U_l^n(t + S_l^n(B_l^n(t_0))) - B_l^n(t_0),\ l = 1, 2, \ldots, L.$$

In condition (C) we would have to require, in addition, that the sequences $\{U^n, n \geq 1\}$ be $C^+([0, \infty), \mathbf{R}^L)$-exponentially tight of order n. In the case of i.i.d. service times this would amount to requiring that the service times have finite exponential moments of every order.

Thus, we have a certain leeway when choosing the function G: we can either restrict ourselves to few equations or write down a more detailed system. It makes sense keeping as few equations as necessary for establishing either ergodicity or non-ergodicity.

Moreover, it is easy to see that we can strenghten the assertions of Theorem 1, Corollary 1, Theorem 2 and Corollary 2 by formulating the conditions in terms of $\mathcal{A}$ rather than $\mathcal{M}$, i.e., by considering "the LD-limit" instead of "the fluid model". This would give us "minimal" over all the fluid models conditions on the fluid trajectories. We have chosen not to state the results in this form because in applications it is easier to check the required properties for the fluid model than for the LD-limit, which is hard to characterise. Note that, in general, the LD-limit is larger than "the fluid limit" (cf., Dai [5], Meyn [12]). According to the next lemma the LD-limit has the properties of self-similarity and compactness, analogous to those of the fluid model.

Lemma 5.1 *The sets $\mathcal{A}(f_0), f_0 \in \mathcal{V}_{b,c,1,L}^+$, have the following properties:*

(self-similarity) for $k = 1, 2, \ldots$ and $f_0 \in \mathcal{V}_{b,c,1,L}^+$,

$$\mathcal{A}(k^{-1} f_0 \circ k) = \{f * k,\ f \in \mathcal{A}(f_0)\},$$

in particular,

$$\mathcal{A}_t\left(\frac{1}{k} f_0 \circ k\right) = \frac{1}{k} \mathcal{A}_{kt}(f_0) \circ k,\ t \geq 0,$$

(compactness) for every $a > 0$ the set $\bigcup_{f_0 \in \mathcal{V}_{b,c,1,L}^+ : \|f_0\| \leq a} \mathcal{A}(f_0)$ is relatively compact in $C([0, \infty), \mathcal{V}_{b,c,1,L}^+)$.

Proof. To check the self-similarity property we note that

$$\frac{1}{k} F^{n(m_n)}(kx, kt) = F^{n(km_n)}(x, t),$$

since both these numbers equal $(nm_n k)^{-1} N^n(\lfloor nm_n kx \rfloor, nm_n k(1 + t))$. The required now follows by the definition of the set $\mathcal{A}$.

Let $K = \{f_0 \in \mathcal{V}_{b,c,1,L}^+ : \|f_0\| \leq a\}$. Since by Lemma 4.2 $\mathcal{A}(f_0) \subset \mathcal{M}(f_0)$, the set $\bigcup\{\mathcal{A}(f_0), f_0 \in K\}$ is relatively compact if the set $\bigcup\{\mathcal{M}(f_0), f_0 \in K\}$ is relatively compact. The latter is implied, in view of the definition of the topology on $\mathcal{V}_b^+$, by the limits

$$\lim_{r\to\infty}\sup_{f_0\in K}\sup_{f\in\mathcal{M}(f_0)}\sup_{\substack{x>r,\\ y>r}}|f_l(x,t)-f_l(y,t)|=0,\ t\geq 0, l=1,...,L,$$

$$\lim_{\delta\to 0}\sup_{f_0\in K}\sup_{f\in\mathcal{M}(f_0)}\sup_{\substack{|t-s|\leq\delta,\\ s,t\leq T}}\sup_{x\geq 0}|f_l(x,t)-f_l(x,s)|=0,\ T>0, l=1,...,L,$$

which follow by (3.3) and (3.6). The lemma is proved.

Note that, since the set $\bigcup\{\mathcal{M}(f_0), f_0\in K\}$ is closed, the above argument also proves the compactness property of the fluid model.

We now discuss some other properties of the fluid model. Let us denote by ρ_m the nominal load at node m:

$$\rho_m=\sum_{l\in c(m)}\frac{\overline{\lambda}_l}{\mu_l}, \tag{5.9}$$

where the $\overline{\lambda}_l$ satisfy the system (we assume that the spectral radius of matrix (p_{kl}) is less than unity):

$$\overline{\lambda}_l=\lambda_l+\sum_{k=1}^{L}\overline{\lambda}_k p_{kl}. \tag{5.10}$$

The interesting nontrivial case is

$$\rho_m<1,\ m=1,...,M, \tag{5.11}$$

since if there exists a node with $\rho_m\geq 1$, then (for a general class of interarrival and service-time distributions) the associated stochastic process is easily non-ergodic.

Let us denote by $\mathbf{0}$ the "empty state" of the fluid model, i.e., the state with no fluid at the nodes. The following observation has been made by A. Stolyar.

Proposition 5.1 *For the FIFO service discipline under the condition (5.11) there exist an initial state f_0 of the fluid model with $\|f_0\|=1$, time t_0 and (fluid) trajectory f (with $f(0)=f_0$) such that $f(t)=\mathbf{0}, t\geq t_0$.*

Intuitively, the state f_0 is chosen so that all the fluid is at the node with the maximum nominal load, the different "fluids" being mixed up at certain proportions. Thus, when all the loads are less than unity, it is never the case for the FIFO discipline that all fluid trajectories for every "non-empty" initial state tend to infinity.

We also note the following simple fact, which is valid for every service discipline.

Proposition 5.2 *If $\rho_m\leq 1$ for all $m\in\{1,...,M\}$, then the trajectory $f(t)=\mathbf{0}, t\geq 0$, is a trajectory of the fluid model.*

Another important property of the fluid model is non-uniqueness of fluid trajectories.

Proposition 5.3 *Let $\rho_m\leq 1, m=1,2,\ldots,M$, and there exist a fluid trajectory $\hat{f}$ such that*

$$\overline{\lim_{t\to\infty}}\frac{\|\hat{f}(t)\|}{t}>0.$$

Then the set $\mathcal{M}(\mathbf{0})$ of fluid trajectories that start at the "empty" initial state contains at least two trajectories.

Indeed, according to Proposition 5.2, there is the fluid trajectory $f(t) = \mathbf{0}, t \geq 0$. On the other hand, let us consider the sequence $f_n = \hat{f} * t_n$, where $t_n \to \infty$ are chosen so that $\|\hat{f}(t_n)\|/t_n \geq c > 0$. According to the self-similarity property, the f_n are also fluid trajectories with initial states $f_n(0) = t_n^{-1}\hat{f}(0) \circ t_n$, which converge to the "empty" initial state as $n \to \infty$. By the compactness property the sequence $\{f_n, n \geq 1\}$ has a limit point $\bar{f}$, which is a fluid trajectory as well. In addition, obviously, $\bar{f}(0) = \mathbf{0}$. Since $\sup_{t\geq 1} \|f_n(t)\|/t = \sup_{t \geq t_n} \|\hat{f}(t)\|/t$, it follows that $\sup_{t\geq 1} \|\bar{f}(t)\|/t \geq c > 0$, which shows that $\bar{f}$ is not the "zero" trajectory. Proposition 5.3 is proved.

Let us note that for a network of one node ($M = 1$) the fluid trajectories are unique so that for all FIFO networks the only initial states that have several outgoing trajectories are states with at least two "empty" nodes. This follows by the fact that for the FIFO discipline during the period of time $\sum_{i \in c(m)} q_i(0)/\mu_i$ the "flow" out of node m is specified uniquely.

Let us now show how one can check the hypotheses of Theorem 1 for basic non-trivial examples of non-ergodic networks. We start with Bramson's example [2], which was the first example of non-recurrence for a network with the FIFO discipline under condition (5.11). This network consists of two nodes, therefore, according to the above remark, there is a unique fluid trajectory for every "non-empty" initial state. This implies, in particular, that for every "non-empty" initial state the fluid trajectory is "a fluid limit". The reasoning below refers to certain properties of the fluid and LD limits, which we do not prove here. They can be derived from the network's properties in Bramson [2] by passing to the limit in the corresponding statements. Theorem 2 in [2] allows us to conclude that for every initial state f_0 such that

$$q_1(0) = H, \sum_{l>1} q_l(0) \leq H/50,$$

there exists a time $T(f_0)$, for which

$$q_1(T(f_0)) \geq 100H, \sum_{l>1} q_l(T(f_0)) \leq H, \; \|f(t)\| > \frac{H}{4}, \; t \in (0, T(f_0)).$$

Let us consider an initial state $\hat{f}_0$ for which $q_1(0) = 1$. Then for a trajectory $\hat{f}(t)$ with $\hat{f}(0) = \hat{f}_0$

$$\underline{\lim}_{t\to\infty} \frac{\|\hat{f}(t)\|}{t} > C > 0,$$

where

$$C = \frac{100}{(4+\delta)T(\hat{f}_0)}$$

(for some $\delta > 0$). Taking the sequence of times $t_1 = T(\hat{f}_0)$, $t_2 = T(\hat{f}(t_1))$, $t_3 = T(\hat{f}(t_2))$, ... we have, for some $c_1 > 0, c_2 > 0$,

$$\underline{\lim}_{k\to\infty} \frac{t_{k+1}}{t_k} > c_1 100 > 1, \quad \overline{\lim}_{k\to\infty} \frac{t_{k+1}}{t_k} < c_2 100 < \infty.$$

The sequence $\{t_k, k = 1, 2, \ldots\}$ plays the role of the sequence 2^k in Theorem 1 (cf., the remark after Corollary 2).

One can show, in addition, that $T(f_0)$ is locally Lipshitz-continuous in a neighbourhood of $\hat{f}_0$:

$$|T(\hat{f}_0) - T(f_0)| \leq c_3 \sup_{x\geq 0} |\hat{f}_0(x) - f_0(x)|$$

for all f_0 from a sufficiently small neighbourhood of $\hat{f}_0$. This observation allows us to check the conditions on the neighbourhood in Theorem 1 (with $R(f, g) = \sup_{x\geq 0} |f(x) - g(x)|$).

Another example is the sequence of networks with a growing number of nodes found by Bramson in [3], where each network is non-ergodic and the nominal loads at the nodes uniformly tend to zero (as the network's index grows). Basically, the construction in [3] is not much different from the one used in the proof of non-recurrence in [2]. The important distinction from the point of view of fluid dynamics, however, is that one has to prove that for all initial states indicated by Bramson the fluid trajectories go to infinity, which does not follow in this case by passing to the limit in the results of [3], since the networks in [3] contain large numbers of nodes and the fluid models might not have the same trajectories as the fluid limits. The idea of the proof is that for these particular instances fluid dynamics decomposes into a number of "switching chains" of fluid nodes, for which fluid trajectories are unique. We omit details.

Finally, in the example found by Rybko and Stolyar [16] for a priority network of two nodes, the fluid dynamics is finite-dimensional and is specified by the four-vector $\{q_l(t), l = 1, \dots, 4\}$. The fluid trajectories are piecewise linear vector-functions and verification of the hypotheses of Theorem 1 is a simplified version of the reasoning used above for Bramson's example.

It is an open question to what extent the hypotheses of Theorem 1 or 2 are necessary for non-ergodicity or non-recurrence of open networks with general structure and service discipline. We make just few comments. The condition on the growth rate of the fluid trajectories appears to be important. Say, in the slightly more general framework of random walks in nonnegative orthants there are instances when an Euler-limit process, which is an analogue of a fluid-limit process, slowly goes to infinity, but the underlying stochastic process is recurrent, see, e.g., Malyshev [11]. We conjecture that if in our setting the total amount of fluid goes to infinity slowly enough, the underlying stochastic process can be null recurrent, but it cannot be positive recurrent. Next, one could conjecture that the ergodicity conditions are insensitive to the service time distributions and are determined only by the mean values. As a supporting argument, note that the fluid trajectories depend only on the mean values and not on the distributions. On the other hand, it could be the case that the fluid-limit behaviour (say, "dissipation probabilities", cf. Malyshev [11]) depends critically on the distributions and that there exist examples, for which conditions of Theorem 1 do not hold, but the trajectories of the fluid limit tend "on average" to infinity for some service-time distributions and tend to zero for others (for the same mean values). All this requires further study.

References

[1] P. Billingsley, Convergence of Probability Measures, Wiley, 1968.

[2] M. Bramson, Instability of FIFO queueing networks, Ann. Appl. Prob. 4 (1994), pp. 414 – 431.

[3] M. Bramson, Instability of FIFO queueing networks with quick service times, Ann. Appl. Prob. 4 (1994), pp. 693 – 718.

[4] M.I. Freidlin and A.D. Wentzell, Fluctuations in Dynamical Systems, Nauka, 1979 (in Russian, English translation: Springer, 1984).

[5] J.G. Dai, On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid models, Ann. Appl. Prob. 5 (1995), pp. 49 – 77.

[6] J.D. Deuschel and D.W. Stroock, Large deviations, Academic Press, New York, 1989.

[7] S.N. Ethier and T.G. Kurtz, Markov Processes: Characterization and Convergence, Wiley, New York, 1986.

[8] P.R. Kumar and T. Seidman, Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems, IEEE Trans. Automat. Control 35 (1990), no.3, pp. 289 - 290.

[9] T. Lindvall, Weak convergence of probability measures and random functions in the function space $D[0,\infty)$, J. Appl. Prob., 10 (1973), pp. 109-121.

[10] R.Sh. Liptser and A.N. Shiryaev, Theory of Martingales, Nauka, 1986 (in Russian, English translation: Reidel, 1989).

[11] A.V. Malyshev, Networks and dynamical systems, Adv. Appl. Prob. 25 (1993), pp. 140 – 175.

[12] S.P. Meyn, Transience of multiclass queueing networks and their fluid models, Ann. Appl. Prob. 5(1995), pp. 946 – 957.

[13] A. Puhalskii, On functional principle of large deviations, in: New Trends in Probability and Statistics, V.Sazonov and T.Shervashidze (eds.), 1991, VSP/Moks'las, Utrecht, v. 1, pp. 198-218.

[14] A. Puhalskii, Large deviation analysis of the single server queue, Queueing Systems 21(1995), pp. 5–66; Erratum: Queueing Systems 23(1996) p. 337.

[15] A. Puhalskii and W. Whitt, Functional large deviation principles for first-passage-time processes, Ann. Appl. Prob. 7 (1997), no. 2, pp. 362–381.

[16] A.N. Rybko and A.L. Stolyar, On ergodicity of random processes describing functioning of open networks, Problems of Information Transmission 28 (1992) pp. 2 – 26. (in Russian).

[17] A.V. Skorohod, Limit theorems for random processes, Theory of Probability and its Applications, 1 (1956) pp. 261 -292 (in Russian).

[18] A. Stolyar, On the stability of multiclass queueing networks: a relaxed sufficient condition via limiting fluid processes, Markov Proc. Rel. Fields, 1 (1995), no. 4, pp. 491 – 512.

[19] S.R.S. Varadhan, Large Deviations and Applications, SIAM, Philadelphia, 1984.