NOLTA 2006
11-14 september
Bologna - ITALY

Queuing Networks with Periodic Fluid Model

Tommaso Addabbo† and Ljupco Kocarev‡

†Information Engineering Department, University of Siena
Via Roma, 56, Siena 53100, Italy
‡Institute for Nonlinear Science, University of California, San Diego
9500 Gilman Drive,La Jolla, CA 92093-0402
Email: addabbo@dii.unisi.it, lkocarev@ucsd.edu

Abstract—In this paper we consider state-dependent Markovian queueing networks and we provide a formal definition of their associated fluid model. We particularly consider queuing networks whose fluid models have global periodic attractors, focusing on the relationships which may occur in such cases between the network stochastic dynamics and the deterministic dynamics of its corresponding fluid model.

1. Introduction

Queuing networks has been successfully utilized for the analysis of several different real-world problems. In the parlance of queueing theory, a queueing system is a place where *customers* arrive to receive service from a service facility (i.e. a *server*). In these models the generic terms "customers" and "servers" may represent various entities depending on the application. In the case of a data switching network, "customers" are packets that arrive at the switching node and "servers" are transmission channels. Recently, queueing theory has been considered for the modeling and understanding of the joint action of genes, proteins and RNA molecules (interwoven in intricate interdependencies commonly known as genetic networks) [1]. In this approach, customers represent molecules of some sort, waiting for chemical reactions.

In this paper we consider *Markovian (or exponential) queueing networks*, for which arrival and service rates, as well as routing probabilities, may depend on queue lengths. Following Mandelbaum and Pats [2], for a such given queuing network we provide a formal definition of its associated fluid model, and we focus on the relationships which may occur between the network stochastic dynamics and the deterministic dynamics of its corresponding fluid model. We give an example of queuing network whose fluid model has a globally attractive limit cycle and as an interesting issue we show that, depending on the network parameters, the fluid model allow for a statistical characterization of the network long-run dynamical behavior. The example is properly chosen to show that the analysis of the fluid model makes sense also when some components of the queuing network periodically become empty (weak load condition).

2. The stochastic model

We consider an n-node open queuing network in which each node is a $M_\xi/M_\xi/1-FIFO$ single station with one server [1]. We denote with $Q_i(t)$ the total number of customers at node i at time t (i.e., the number of customers in service plus the number of customers waiting in the buffer i), while λ_i and μ_i denote the exogenous arrival rate and service rate for the node i, respectively. After being served at node i, a customer can randomly either leave the system or instead he can join any queue. In the latter case, the customer who is just leaving node i will join the queue at node j depending on the transition probability P_{ij} which characterizes the customer flow from node i to node j. The $n \times n$ square matrix $\mathbf{P} = (P_{ij})$ is the *transition probability matrix* of the queuing network, where $0 \le P_{ij} \le 1$ and $\sum_{j=1}^{n} P_{ij} \le 1$. In this paper we assume that both arrival rates and service rates as well as the transition matrix are locally Lipschitz continuous functions of the state $\mathbf{Q}$. In this case, by denoting with $\mathbb{N}$ the set of natural numbers, the dynamic evolution of the queuing network is the realization of a $\mathbb{N}^n$-valued stochastic process $\mathbf{Q} = \{\mathbf{Q}(t), t \ge 0\}$ that satisfies a relation of the form

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \mathbf{A}(t) + \mathbf{F}(t) - \mathbf{D}(t), \qquad t \ge 0, \quad (1)$$

where the column vector $\mathbf{Q}(0)$ is the initial state of the queuing network, $\mathbf{A}(\cdot)$, $\mathbf{F}(\cdot)$, $\mathbf{D}(\cdot)$ are n-dimensional counting processes related to the customer arrivals from the outside of the network (exogenous arrivals), to the customer flows inside the network (endogenous arrivals), and to the service completions (departures), respectively. For the i-th component in (1) we assume

$$A_i(t) = N_i^+\left(\int_0^t \lambda_i\left(\mathbf{Q}(s)\right) ds\right) \quad (2)$$

$$D_i(t) = N_i^-\left(\int_0^t \mu_i\left(\mathbf{Q}(s)\right) 1_{(0,\infty)}(Q_i(s)) ds\right) \quad (3)$$

$$F_i(t) = \sum_{j=1}^{n} \int_0^t 1_{\pi_{ji}(s-)}\left(U_j(D_j(s)\right) dD_j(s) \quad (4)$$

where in general the writing $Q_i(s-)$ denotes the left limit of $Q_i(\cdot)$ at s, assuming $Q_i(0-) = Q_i(0)$. In the above equations the used notation has the following meaning: for

$i = 1, \ldots, n$, the quantities $N_i^+(\cdot), N_i^-(\cdot)$ are mutually independent standard Poisson processes (with rate 1). Given any set B, function $1_B : B \to \{0, 1\}$ is an *indicator function* defined as $1_B(x) = 1$ if $x \in B$, or $1_B(x) = 0$ otherwise. In (3) the set $(0, \infty)$ is the set of strictly positive reals and the indicator function $1_{(0,\infty)}(\cdot)$ is used since no customer departure from a station can occur while this station is empty. For $j = 1, \ldots, n$, the sequences $\{U_j(l)\}_{l\in\mathbb{N}}$ are assumed to be sequences of i.i.d. random variables uniformly distributed on $[0, 1)$, and for $i, j = 1, \ldots, n$ and $s \geq 0$, the set $\pi_{ji}(s-)$ is equal to $\pi_{ji}(s-) = \left\{x \in [0, 1) : \sum_{k=1}^{i-1} P_{jk}(\mathbf{Q}(s-)) \leq x < \sum_{k=1}^{i} P_{jk}(\mathbf{Q}(s-))\right\}$, where for $j, k = 1, \ldots, n$, the terms P_{jk} are elements of the transition matrix function $\mathbf{P}$. In this expression, the left limit is used since, when a customer is served at time t_0, she/he is instantaneously routed according to the transition matrix $\mathbf{P}(\mathbf{Q}(t_0-))$ which refers to the networks state at $\mathbf{Q}(t_0-)$. Indeed, the transition matrix $\mathbf{P}(\mathbf{Q}(t_0))$ is defined only when the customer served at t_0 has been definitely routed somewhere. The introduced above quantities have the following interpretations: for $t \geq 0$, $A_i(t)$ and $F_i(t)$ represent the cumulative number of exogenous and endogenous arrivals respectively to station i during $[0, t]$; when a customer at time t is served at node j, the quantity $D_j(t)$ increases by 1 from $D_j(t-)$. If the random variable $U_j(D_j(t)) \in \pi_{jk}(t-)$ then the customer departing form node j is routed to node k. Note that for each $t \geq 0$, $\bigcup_{i=1}^{n} \pi_{ji}(t) \subseteq [0, 1)$.

It is worth enphasizing that we assume the processes N_i^+ and N_i^- and the sequences $\{U_i(l)\}_{l\in\mathbb{N}}$ all mutually independent. Moreover we assume that the spectral radius, $r(\mathbf{P}(\cdot))$, of the transition matrix $\mathbf{P}$ is such that $r(\mathbf{P}(\mathbf{v})) < 1$ for all $\mathbf{v} \in \mathbb{R}_+^n$ and that the two vector-valued functions $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ satisfy a linear growth constraint, that is there exists a positive constant Δ such that for all $\mathbf{v} \in \mathbb{R}_+^n$ it results $\max\{\|\boldsymbol{\lambda}(\mathbf{v})\|, \|\boldsymbol{\mu}(\mathbf{v})\|\} \leq \Delta(1 + \|\mathbf{v}\|)$. As discussed in [2], since the queuing network process described above is a Markov jump process on $\mathbb{N}^n$, this latter assumption represents a sufficient condition for the non-explosion of the network's state $\mathbf{Q}$, that is $P\{\|\mathbf{Q}(t)\| < \infty, t \geq 0\} = 1$. Almost surely, the paths of the stochastic process $\{\mathbf{Q}(t), t \geq 0\}$ consist of n piecewise constant right-continuous non-negative integer-valued functions with finite left limits. Moreover, for $i = 1, \ldots, n$, stochastic processes $\{A_i(t), t \geq 0\}, \{D_i(t), t \geq 0\}$ can be referred to as *time-changed Poisson processes*.

3. Fluid Models and Fluid Approximations

Fluid approximations are formal deterministic approximations to queuing networks. Frequently these are obtained as scaling limits, where arrival and service rates are speeded up, while attenuating the stochastic fluctuations by a same quantity. The limit for this procedure suppresses the random fluctuations around what appears to be the main trajectory of the process. From a theoretical point of view, a fluid model of a queuing network is obtained as the limit of a sequence of stochastic processes $\{\mathbf{q}^k\} = \{\mathbf{q}^k(t), t \geq 0\}$ which is determined by means of simple transformations of parameters of the starting stochastic system [1]. Here, the superscript k indicates that the corresponding quantity is related to the k-th network of such a sequence. Accordingly, let us consider a queuing network with initial condition $\mathbf{Q}_0$ and whose stochastic evolution is described by a random process $\mathbf{Q}$ satisfying relationship (1) and arising assumptions. We introduce with $k = 1, 2, \ldots$ a sequence of stochastic processes $\{\mathbf{q}^k\}$, all of them having the same initial condition $\mathbf{Q}_0$, such that

$$\mathbf{q}_0^k = \mathbf{Q}_0 \tag{5}$$

$$\boldsymbol{\lambda}^k(\mathbf{v}) = k\boldsymbol{\lambda}(\mathbf{v}) \qquad \boldsymbol{\mu}^k(\mathbf{v}) = k\boldsymbol{\mu}(\mathbf{v}) \qquad \mathbf{P}^k(\mathbf{v}) = \mathbf{P}(\mathbf{v}) \tag{6}$$

where $\mathbf{v} \in \mathbb{R}^n$ and such that for each $k \in \mathbb{N}$ the stochastic process $\mathbf{q}^k$ satisfies relationship (1) and subsequent assumptions, i.e.

$$\mathbf{q}^k(t) = \mathbf{Q}_0 + \mathbf{A}^k(t) + \mathbf{F}^k(t) - \mathbf{D}^k(t) \qquad t \geq 0 \tag{7}$$

where, for the i-th component in (7) we have

$$A_i^k(t) = \frac{1}{k} N_i^+\left(k \int_0^t \lambda_i\left(\mathbf{q}^k(s)\right) ds\right) \tag{8}$$

$$D_i^k(t) = \frac{1}{k} N_i^-\left(k \int_0^t \mu_i\left(\mathbf{q}^k(s)\right) 1_{(0,\infty)}(q_i^k(s)) ds\right) \tag{9}$$

$$F_i^k(t) = \sum_{j=1}^{n} \int_0^t 1_{\pi_{ji}^k(s-)}\left(U_j(kD_j^k(s)\right) dD_j^k(s) \tag{10}$$

By construction, in the Markovian stochastic process described by $\mathbf{q}^k$ an arrival event in node i is recorded by increasing the length of the i^{th} corresponding *rational-valued* "queue" by the quantity $1/k$, and the same quantity is subtracted when a departure occurs. Accordingly, with reference to the queuing network described by $\mathbf{Q}$, the process $\mathbf{q}^k$ has arrival and service rates speeded up k times to compensate for the k-times reduced perturbation effects on the state by the stochastic fluctuations. Note, also, that if $k = 1$ then $\mathbf{q}^1 = \mathbf{Q}$, i.e. the stochastic process $\mathbf{Q}$ describing the original queuing network agrees with the first element of the sequence $\{\mathbf{q}^k\}$ of rescaled stochastic processes.

For the dynamical behavior of the asymptotic process $\mathbf{q}^k$ when $k \to \infty$ the following theorem holds.

Theorem 1 (Fluid Limit [2]) *The sequence of stochastic processes $\{\mathbf{q}^k(t), t \geq 0\}$ defined in (5) - (10) for $k \to \infty$ converges in probability uniformly over compacts in $[0, \infty)$ to the unique non-negative absolutely continuous function $\mathbf{q} : \mathbb{R}_+ \to \mathbb{R}_+^n$ which is for $t \geq 0$ the solution of*

$$\mathbf{q}(t) = \mathbf{Q}_0 + \int_0^t \left[\boldsymbol{\lambda}(\mathbf{q}(s)) + \left[\mathbf{P}'(\mathbf{q}(s)) - \mathbf{I}\right]\boldsymbol{\mu}(\mathbf{q}(s))\right] ds + \mathbf{r}(t), \tag{11}$$

where the term $\mathbf{r}(t)$ *takes into account the cumulative effect of the system's state-dependent reflection on the boundaries*

$$\mathbf{r}(t) = \int_0^t \left[\mathbf{I} - \mathbf{P}'\left(\mathbf{q}(s)\right)\right] d\mathbf{y}(s), \quad t \geq 0, \tag{12}$$

where the function $\mathbf{y} : \mathbb{R}_+ \to \mathbb{R}_+^n$ *is continuous, non-negative, and non-decreasing components, such that* $\mathbf{y}(0) = \mathbf{0}$ *and*

$$\int_0^\infty \mathbf{I}_{(0,\infty)}(\mathbf{q}(t)) d\mathbf{y}(t) = \mathbf{0}. \tag{13}$$

The unique solution $\mathbf{q}$ of (11) is referred to as the *fluid limit* of the sequence of processes $\{\mathbf{q}^k(t), t \geq 0\}$ derived from the stochastic process $\{\mathbf{Q}(t), t \geq 0\}$ with initial condition $\mathbf{Q}_0$. By condition (13), since $\mathbf{y}$ is non-negative and due to the continuity of $\mathbf{q}$, it turns out that if the initial condition $\mathbf{Q}_0$ does not have any components on the boundaries, then there exists an interval $[0, t_s]$ for which $\mathbf{r}(t) = \mathbf{0}\ \forall t \in [0, t_s]$. Under this hypothesis, the fluid limit $\mathbf{q}$ over $[0, t_s]$ agrees with the unique solution of the system of ordinary differential equations (ODEs)

$$\frac{d\mathbf{q}}{dt} = \boldsymbol{\lambda}(\mathbf{q}) + \left[\mathbf{P}'(\mathbf{q}) - \mathbf{I}\right]\boldsymbol{\mu}(\mathbf{q}) \equiv \mathbf{h}(\mathbf{q}) \tag{14}$$

when assuming $\mathbf{q}(0) = \mathbf{Q}_0$. In general, when taking into account the effects of the boundaries, instead of (14) the following equation must be considered [2]:

$$\frac{d\mathbf{q}}{dt} = \mathbf{h}(\mathbf{q}) + [\mathbf{I} - \mathbf{P}'(\mathbf{q})]\tilde{\mathbf{m}}(\mathbf{q}) \equiv \mathbf{H}(\mathbf{q}), \tag{15}$$

where the function $\mathbf{h}$ is defined in (14) and function $\tilde{\mathbf{m}}$ is unambiguously determined by the state-dependent reflection, as described in the following. Referring to the introduced notation, let denote by $\mathcal{I}(t) \subseteq \{1, 2, \ldots, n\}$ the subset of indices i such that $h_i(\mathbf{q}) < 0$ and such that the corresponding state's component q_i is zero, i.e. we have $i \in \mathcal{I}(\mathbf{q}) \Leftrightarrow q_i = 0, h_i(\mathbf{q}) < 0$, for $i = 1, \ldots, n$. Accordingly, the components of $\tilde{\mathbf{m}}(\mathbf{q})$ can be obtained by solving the linear system:

$$\begin{cases} h_i(\mathbf{q}) + \tilde{m}_i(\mathbf{q}) - \sum_{j \in \mathcal{I}(\mathbf{q})} P_{ji}(\mathbf{q})\tilde{m}_j(\mathbf{q}) = 0 & \text{if } i \in \mathcal{I}(\mathbf{q}) \\ \tilde{m}_i(\mathbf{q}) = 0 & \text{otherwise} \end{cases} \tag{16}$$

Expression (15) is referred to as the *fluid model equation* for the queuing network with queue length process $\mathbf{Q}$.

3.1. A Periodic Fluid Model Example

In a state dependent queuing network, *since no strong constraints are imposed on the resulting functional form of the deterministic fluid model*, the resulting dynamical system (15) can display a wide range of dynamics, including steady states, periodic behavior and chaos. To give an example, let us consider a simple two-station $M_\xi/M_\xi/1 - FIFO$ network with parameters

$$\lambda_x(\mathbf{Q}^E) = \begin{cases} 6E, & \text{if } 0 \leq Q_x^E, Q_y^E \leq E \\ 6E + 20 f_x^E, & \text{if } Q_x^E > E \\ 6E + \frac{5}{E} f_x^E f_y^E, & \text{otherwise,} \end{cases} \tag{17}$$

$$\mu_x(\mathbf{Q}^E) = \\ = \begin{cases} 3E, & \text{if } 0 \leq Q_x^E, Q_y^E \leq E \\ 3E + 20 f_x^{E^+} + 4 f_y^E, & \text{if } Q_y^E > E \\ 3E + 20 f_x^E, & \text{otherwise,} \end{cases} \tag{18}$$

$$\lambda_y = 0, \qquad \mu_y = 5E, \qquad P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{19}$$

where $E > 0$, $f_x^E = Q_x^E - E$, $f_y^E = Q_y^E - E$ and, for any real number a, we have $(a)^+ = \max\{0, a\}$. By changing parameter E we change the *way the network reacts to its workload*, and for each $E \in \mathbb{R}^+$ the stochastic process $\mathbf{Q}^E$, describing the queuing network with parameters (17)-(19), satisfies all assumptions discussed in Section 2. Therefore, referring to (15), the fluid model of $\mathbf{Q}^E$ has the form $\frac{d\mathbf{q}}{dt} = \mathbf{h}^E(\mathbf{q}) + [\mathbf{I} - \mathbf{P}']\tilde{\mathbf{m}}^E(\mathbf{q}) \equiv \mathbf{H}^E(\mathbf{q})$, where the function $\mathbf{h}^E = (h_x^E, h_y^E)$ is given by

$$h_x^E(\mathbf{q}) = \begin{cases} 3E, & \text{if } 0 \leq q_y \leq E \\ 3E - 4g_y, & \text{if } q_x, q_y > E \\ 3E + \frac{5}{E} g_x g_y - 4g_y, & \text{otherwise,} \end{cases} \tag{20}$$

$$h_y^E(\mathbf{q}) = \begin{cases} -2E, & \text{if } 0 \leq q_x, q_y \leq E \\ -2E + 20(g_x)^+ + 4g_y, & \text{if } q_y > E \\ -2E + 20g_x, & \text{otherwise,} \end{cases} \tag{21}$$

where $g_x = q_x - E$ and $g_y = q_y - E$. The analysis of (20)-(21) reveals that there are points on the boundary ($q_x = 0$) in which h_x^E is negative for all $E > 0$, and that there are some points on the boundary ($q_y = 0$) in which h_y^E is negative for all $E > 0$. Indeed it results $h_x^E(0, q_y) < 0$ if $q_y > 4E/3$ and $h_y^E(q_x, 0) < 0$ if $0 \leq q_x < 11E/10$. Accordingly, solving system (16), function $\tilde{\mathbf{m}}^E$ has components

$$\tilde{m}_x^E(\mathbf{q}) = \begin{cases} -h_x(\mathbf{q}) = 9q_y - 12E, & \text{if } q_x = 0, q_y > \frac{4E}{3} \\ 0, & \text{otherwise,} \end{cases} \tag{22}$$

$$\tilde{m}_y^E(\mathbf{q}) = \begin{cases} -h_y(\mathbf{q}) = 22E - 20q_x, & \text{if } 0 \leq q_x \leq \frac{11E}{10}, q_y = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The dynamical system (15) with $\mathbf{h}^E$ and $\tilde{\mathbf{m}}^E$ defined as in (20)-(21) and (22)-(23) has the global periodic attractor plotted in solid line in Fig. 1.b for $E = 500$. A rigorous analysis of the global attractiveness of the limit cycle is here omitted, however it is not difficult to verify that for this system every trajectory must eventually hit the boundary $q_x = 0$ in $q_y > 4E/3$ entering in a finite time the periodic cycle after a transient evolution flattened on the q_y axis. In the same figure, two stochastic dynamics of two queuing networks with primitives (17)-(19) for $E = 50$ and 500 respectively are plotted.

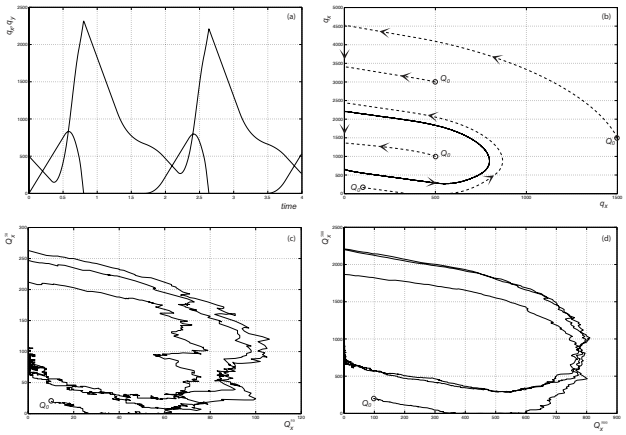


Figure 1: *(a)*: dynamical evolution of (q_x, q_y) for the fluid trajectory over the limit cycle ($E = 500$); *(b)*: the global attractiveness of the limit cycle is here shown for several initial conditions ($E = 500$). *(c)* and *(d)*: stochastic dynamics of two queuing networks with primitives (17)-(19) for $E = 50$ and $E = 500$ respectively.

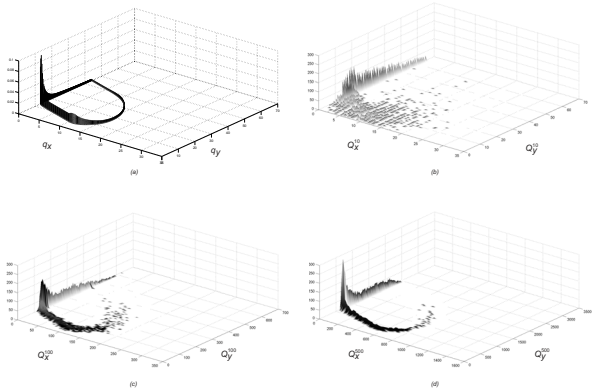


Figure 2: *(a)* - The unique invaiant distribution over the limit cycle support φ for the fluid model of the queuing network (17)-(19) with $E = 10$. *(b), (c) and (d)* - Distributions of 10000 realizations evaluated at $t = 10^4$s of the stochastic process $\mathbf{Q}^E(\cdot)$ for $E = 10, 100$ and 500.

4. Periodic Queuing Networks: Long-Run Behavior

In this section we analyze the stochastic dynamics from a probabilistic point of view. In particular, we show that for queuing networks with periodic fluid models the system state under the stochastic evolution approaches a stationary invariant distribution which is intimately related to the unique invariant probability measure existing for the periodic fluid model [1]. To get insight this point, let write (15) as a dynamical system described by $\frac{d\mathbf{q}}{dt} = \mathbf{H}(\mathbf{q})$, and let assume the existance of a globally attractive limit cycle with support $\varphi \subset \mathbb{R}^n$. By denoting with $\mathcal{B}(\mathbb{R}^n)$ the standard Borel σ-algebra over $\mathbb{R}^n$, it can be proved [1] that the function $\Pi : \mathcal{B}(\mathbb{R}^n) \to [0, 1]$ defined by the curvilinear integral $\Pi(A) = \int_{A\cap\varphi} \frac{dl}{T\cdot\|\mathbf{H}(\mathbf{q})\|}$ is a probability measure over $\mathbb{R}^n$. The following property holds [1]

Theorem 2 *The measure* Π *is the unique invariant probability measure over* $\mathbb{R}^n$ *with respect to the periodic dynamical system.*

In eventually periodic systems the stationary distribution related to the probability measure Π is impulsive over the zero-measure support of the limit cycle. In such case, the unique stationary (impulsive) distribution is described over φ by a curvilinear probability density function (pdf) $\psi : \varphi \to \mathbb{R}_+$ defined as $\psi(\mathbf{q}) = \frac{1}{Tv(\mathbf{q})}$.

We now provide some numerical evidence that the state distribution of a queuing network with periodic fluid model approaches under the evolution a stationary invariant distribution strongly related to the unique impulsive stationary distribution of the periodic fluid model. Referring to our example, this issue appears clear in Fig. 2. The numerical analysis of the stochastic evolution showed that, regardless of the the initial distribution, a crater-shaped stationary distribution is approached, being the ridge of the crater positioned over the support φ of the fluid limit cycle. This result can be explained observing that the dynamics in queuing networks evolves under the strong influence of the underlying fluid deterministic velocity field. Since in our case the fluid system has a globally attractive limit cycle, over the periodic trajectory the deterministic state variable which corresponds to the zero Lyapunov exponent, the phase variable, is neutrally stable and hence, depending on the system parameters, the dominant effect of the stochastic perturbations is a phase diffusion, and around φ a distribution proportional to the natural invariant curvilinear pdf ψ is consequently approached.

In conclusion, the dynamical analysis of the fluid model can provide useful information for *predicting* the evolution of a queuing network, especially when its fluid model has a globally attractive limit cycle. The example has shown that the queuing network dynamics can be strongly influenced by the deterministic velocity field even if the queues become empty (i.e. even if the network has a small workload). This is in sharp contrast to the intuitive idea that fluid approximations can be sensibly taken into account only for queuing networks with overloaded servers.

References

[1] T. Addabbo, L. Kocarev and R. J. Williams, "Periodic Dynamics in State Dependent Queuing Networks" *to be submitted to Physica A*.

[2] A. Mandelbaum, G. Pats, "State-Dependent Stochastic Networks. Part I: Approximations and Applications with Continuous Diffusion Limits", *The Annals of Applied Probability*, Vol.8, No.2, 1998, pp.569–646.