The Annals of Applied Probability
2003, Vol. 13, No. 4, 1423–1448

LARGE-DEVIATIONS ANALYSIS OF THE FLUID APPROXIMATION FOR A CONTROLLABLE TANDEM QUEUE

BY ALEXANDER GAJRAT,[1] ARIE HORDIJK AND AD RIDDER

Vrije Universteit Amsterdam, Leiden University and Vrije Universteit Amsterdam

A fluid approximation gives the main term in the asymptotic expression of the value function for a controllable stochastic network. The policies that have the same asymptotic of their value functions as the value function of the optimal policy are called asymptotically optimal policies. We consider the problem of finding from this set of asymptotically optimal policies a best one in the sense that the next term of its asymptotic expression is minimal. The analysis of this problem is closely connected with large-deviations problems for a random walk.

1. Introduction. Many problems of optimal control in stochastic queueing networks or, more generally, in random walks are difficult to solve explicitly or numerically. One of the reasons may be the large state space that usually is involved when one applies techniques from Markov decision theory such as policy improvement or value iteration. However, considering these stochastic models without control, deterministic models have been developed that approximate them in some asymptotical sense. A well-known approximative model is the fluid model. In recent years, this technique of fluid approximation has been applied to stochastic control problems as well; see, for example, Atkins and Chen (1995), Avram, Bertsimas and Ricard (1995), Bäuerle and Rieder (2000), Gajrat, Hordijk, Malyshew and Spieksma (1997), Gajrat and Hordijk (2000), Maglaras (1999), Maglaras (2000) and Weiss (1999). This approach leads to a so-called fluid control model or problem.

Generally, the construction of an optimal solution in the fluid control model is again a difficult task, but some solution methods for specific problems have been derived, for instance, scheduling problems [Atkins and Chen (1995), Avram, Bertsimas and Ricard (1995) and Weiss (1995, 1999)] and for service control in queueing networks [Bäuerle and Rieder (2000) and Gajrat and Hordijk (1999, 2000a)]. Also conditions on the existence of optimal fluid controls have been studied [Pullan (1995, 1996)]. Having solved the deterministic fluid model, the next step is to construct a control or policy in the stochastic system in such a way that its asymptotic behavior corresponds to the optimal solution of the

---

Received January 2001; revised September 2002.

[1]Supported by the project "Stochastic Networks" of the Netherlands Organization for Scientific Research NWO.

AMS 2000 subject classifications. Primary 60K25, 93E20; secondary 60F10.

Key words and phrases. Stochastic network, fluid optimal control problem, asymptotic optimality, large deviations.

deterministic problem. These policies are called asymptotically fluid optimal (a.f.o.) policies. The translation problem has been studied in the context of the heavy-traffic approximation of the control of stochastic networks; see Harrison (1996, 1998), Kelly and Laws (1993), Maglaras (2002) and Bell and Williams (2001). In Harrison (1996) a general method, called BIGSTEP, is developed. In Harrison (1998) it is shown for a two-station model that this method leads to asymptotically optimal policies. The discrete-review policies and tracking policies in Maglaras (1999, 2000) use similar methods for the fluid approximation.

As soon as one can find a.f.o. policies for a given stochastic system, the problem is solved from a fluid point of view. However, there are many types of a.f.o. policies, and some have the same structure and others are different. For example, the policies may be defined by different switching curves [Gajrat, Hordijk, Malyshev and Spieksma (1997) and Gajrat and Hordijk (2000a)], discrete-review policies [Maglaras (1999)] or tracking-policies [Bäuerle (2000) and Maglaras (2000)]. So the natural question then is: can we find in this set of a.f.o. policies some policies that are "better" than others, where better means that they dominate other policies in asymptotic behavior. In this paper, we study and answer this question for a particular model.

We will consider the fluid approximation of a controllable stochastic tandem queue, discrete in time and space. The fluid approximation is continuous in time and space and is used to get the first term of the asymptotic of the value function of the optimal policy. More precisely, for a linear cost function, the value function associated with the discrete decision rule a has the form $V_a(x_N) = N^2 F_u(x) + o(N^2)$; see Gajrat and Hordijk (2000a). Here, N is a scaling parameter such that $x_N/N \to x$ and $F_u(x)$ is the value function of the corresponding fluid queue associated with the continuous control u. Specifically, this asymptotic holds for the optimal discrete value function $V_{\mathrm{opt}}(x_N)$ and the optimal fluid value function $F_{\mathrm{opt}}(x)$. The same asymptotic appears for a class of a.f.o. policies that are characterized by switching curves separating regions with different actions. Different switching curves give the same first term of the asymptotic of the value function (see Remark 2).

So we can reformulate the question: Which switching curve gives the smallest next term in the asymptotic and what is the order of this next term? For our model, the natural choice (see Remark 5) of the switching curve is given by the function $h(x) = [\gamma \ln x]$. We will show that there are two main types of asymptotics for the value function and that these types depend on either γ being greater or less than some constant.

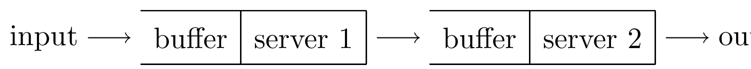
Threshold strategies of this type have been considered in Kelly and Laws (1993) and Harrison (1998) for different two-station models. The threshold or safety stock they use in the heavy-traffic approximation is $r(N) = \gamma \ln N$, where N is the time-scale parameter. If we realize that the fluid optimal control is the optimal policy for a large initial state $x_N \simeq Nx$, we find an interesting similarity in the shape of the asymptotically optimal policy for the heavy-traffic and the fluid approximations. The model studied in Bell and Williams (2001) is a two-server parallel queue with control of assigning jobs to available servers. Again, a threshold policy is asymptotically optimal in the heavy-traffic limit.

While in our paper an a.f.o. policy based on a switching curve is constructed for a specific tandem network, methods for constructing an a.f.o. policy for a general stochastic network have been derived in Bäuerle (2000) (tracking policies) and Maglaras (1999) (discrete-review policies). The advantage of the approaches in Bäuerle (2000) and Maglaras (1999) is that they do not demand a detailed description of the space structure for an optimal solution of the fluid model. For the construction of a switching-curve a.f.o. policy, one needs a more detailed geometrical description of the optimal solution [explained in Gajrat and Hordijk (2000b)]. On the other hand, the performance of a switching-curve a.f.o. policy might be "better" (see Remark 4).

Concluding, for our specific controlled queueing network (the tandem queue), a number of switching curves exists, each one yielding an asymptotically fluid optimal policy, but the specific logarithmic curve gives a better second-order asymptotic. We have considered generalizing such a result to other queueing networks and found basically two problems when we tried to extend our techniques. The first problem is that the geometrical description of the space structure of an optimal solution of the fluid model is much more complicated; see Gajrat and Hordijk (2000b). In general, it looks like a stratification of the multidimensional (the dimension is the number of buffers in a network) octant in a set of cones of different dimensions. And inside each cone the policy is homogeneous (one action is defined). The second problem is more probabilistic. Even when we have a clear geometrical picture of the optimal fluid policy, we should define a corresponding stochastic network. To analyze the asymptotical properties of this stochastic network, we should have a theory of large deviations for random walks in multidimensional space with different regions of homogeneity. There are some results [Blinoskii and Dobrushin (1994), Ignatyuk, Malyshew and Scherbakov (1994) and Ignatyuk (1998)] in this direction but they should be extended. Thus, a nontrivial generalization can appear only from developing a corresponding large-deviations technique.

The paper is organized as follows. In Section 2, we describe the control problem and we review the first-order fluid asymptotic. Section 3 states the main theorem concerning the next term of the asymptotic and compares the switching-curve a.f.o. policy with the tracking a.f.o. policy. Section 4 contains the proof of the theorem, which turns out to be quite elaborate, and therefore we choose to split up the proof in a number of lemmas distributed in several sections. The core lies in Sections 4.6 and 4.8, where we deal with the asymptotics for boundary probabilities and where we apply large deviations for random walks to obtain these asymptotics. Finally, Section 5 concludes with a numerical example.

2. Description of the controllable tandem queue. We consider a discrete-time tandem queueing network of two single-server queues, Bernoulli (λ) arrivals, Bernoulli (μ_1 and μ_2) servers and infinite buffers:



The state of the network at time n is $\xi^n = \{(\xi_1^n, \xi_2^n)\}_{n=0,1,2,3,\dots} \in \mathbb{Z}_+^2$, where ξ_i^n is the number of customers in buffer i. At each moment of time n, a server i can choose to:

(i) serve a customer from its buffer (if the buffer is nonempty): $a_i^n = 1$, or
(ii) be idle: $a_i^n = 0$.

The control variables a_i^n denote the state of the server i; $a_i^n = 1$ when the server is serving and $a_i^n = 0$ when it is idle. When the server i is serving a customer, this customer will be removed from buffer i with probability μ_i. With probability λ, a new customer arrives at buffer 1. When a customer is removed from server 1, it moves to the buffer of the second server. When a customer is removed from the buffer of the second server, this customer leaves the network.

So the control $a = \{a^n\}_{n=0,1,\dots}$ defines a Markov chain $\{\xi^n\}_{n=0,1,\dots}$ with dynamics

$$\begin{aligned} \xi_1^{n+1} - \xi_1^n &= \eta_0^n - a_1^n \eta_1^n, \\ \xi_2^{n+1} - \xi_2^n &= a_1^n \eta_1^n - a_2^n \eta_2^n, \end{aligned} \tag{1}$$

where $\{\eta_i^n\}_{n=0,1,\dots}$, $i = 0, 1, 2$, are i.i.d. Bernoulli processes on $\{0, 1\}$ with $E\eta_0^n = \lambda$, $E\eta_i^n = \mu_i$, $i = 1, 2$. The three processes are mutually independent.

We are interested in the following optimal control problem for this class of network. Let T_N be some finite time, $a = \{a^n\}$ a control and $x(N) \in \mathbb{Z}_+^2$ a state, such that $T_N \to \infty$, $\|x(N)\| \to \infty$. Then the value function of the process under control a is defined as

$$V_a(x(N)) := E_{x(N)}^a \sum_{n=0}^{T_N} \xi^n c = E_{x(N)}^a \sum_{n=0}^{T_N} (\xi_1^n c_1 + \xi_2^n c_2), \tag{2}$$

where $E_{x(N)}^a(\cdot)$ denotes the expectation given initial state $x(N)$ and control $a = \{a^n\}$ and where $\xi^n c$ is the inner product of the two vectors: $\xi^n c = \xi_1^n c_1 + \xi_2^n c_2$. The discrete optimal control problem is

$$\min_a V_a(x(N)). \tag{3}$$

By $V_{\text{opt}}(\cdot)$, we denote the value function of an optimal control of this problem. (There exists an optimal control since we are dealing with a finite-horizon problem.) The function V_{opt} cannot be found precisely, but we can try to find an asymptotic of this function or we can formulate more simple problems. We shall investigate an asymptotic expression of the optimal value function. First, we formulate the corresponding fluid control model of the controllable network.

2.1. *The fluid controllable network.* Let $(x^s)_{s\geq 0} = (x_1^s, x_2^s)_{s\geq 0}$ be a continuous deterministic process on $\mathbb{R}_+^2$ with derivatives $\dot{x}^s$ satisfying

$$\begin{aligned}\dot{x}_1^s &= \lambda - \mu_1 u_1^s,\\ \dot{x}_2^s &= \mu_1 u_1^s - \mu_2 u_2^s,\\ 0 &\leq u_s^i \leq 1.\end{aligned}$$

Here $u = (u^s)_{s\geq 0} = (u_1^s, u_2^s)_{s\geq 0}$ is a control that regulates continuously the contents of the fluid buffers. Note that the process (x^s) is determined by control u. Let t be a finite time, u a control and $x \in \mathbb{R}_+^2$. Then the value function under this control is

$$F_u(x) := \int_0^t x^s c\, ds = \int_0^t (x_1^s c_1 + x_2^s c_2)\, ds \qquad \text{with } x^0 = x.$$

The fluid optimal control problem is

$$\min_u F_u(x).$$

By $F_{\text{opt}}(\cdot)$, we denote the value function of the optimal control of this problem, and by (x_{opt}^s), the fluid process or trajectory under the optimal control. We solve the optimal control problem for the following set of parameters:

$$\mu_1 > \mu_2 > \lambda > 0, \qquad c_2 > c_1 > 0. \tag{4}$$

REMARK 1. One can consider also another set of parameters, but for the purpose of this paper this set of parameters is the most significant one, because it is the case where one should introduce a nonlinear switching curve.

The optimal solution for the fluid network with the set of parameters (4) is the following:

$$\begin{aligned}&\text{while } x_1^s > 0 \text{ and } x_2^s > 0: && u_1^s = 0,\ u_2^s = 1,\\ &\text{while } x_1^s > 0 \text{ and } x_2^s = 0: && u_1^s = \mu_2/\mu_1,\ u_2^s = 1,\\ &\text{while } x_1^s = 0 \text{ and } x_2^s = 0: && u_1^s = \lambda/\mu_1,\ u_2^s = \lambda/\mu_2.\end{aligned}$$

Equivalently, the optimal trajectory satisfies the differential equations

$$\dot{x}_{\text{opt}}^s = \begin{cases} (\lambda, -\mu_2), & \text{if } x_1^s > 0 \text{ and } x_2^s > 0,\\ (\lambda - \mu_2, 0), & \text{if } x_1^s > 0 \text{ and } x_2^s = 0,\\ (0, 0), & \text{if } x_1^s = 0 \text{ and } x_2^s = 0.\end{cases} \tag{5}$$

In words: suppose (i) that the initial point x^0 has positive buffers ($x_1^0 > 0$ and $x_2^0 > 0$). Then the optimal trajectory empties buffer 2 at a speed of μ_2 while filling buffer 1 at a speed of λ. There is no flow from buffer 1 to buffer 2. This happens until the second buffer is empty. Now, suppose (ii) that the initial point x^0 lies on the boundary $x_2^0 = 0$, that is, buffer 2 is empty. Then the optimal trajectory empties buffer 1 at a speed of $\mu_2 - \lambda$ while keeping buffer 2 empty by balancing the inflow and outflow (of buffer 2). This happens until buffer 1 is empty as well, whereafter the two buffers remain empty by serving at a speed of the inflow rate λ. Only part (ii) of the optimal trajectory is of interest in our analysis. So, assuming the initial point x^0 has $x_1^0 > 0$ and $x_2^0 = 0$, the optimal trajectory satisfies

$$x_{\text{opt}}^s = x^0 + s(\lambda - \mu_2, 0),$$

as long as $s < x_1^0/(\mu_2 - \lambda)$. The optimal value function becomes

$$F_{\text{opt}}(x^0) = \int_0^t c_1\big(x_1^0 + s(\lambda - \mu_2)\big)\,ds \tag{6}$$

for $t < x_1^0/(\mu_2 - \lambda)$.

2.2. *Asymptotics.* We shall consider the following version of the value function (2) in the original optimal control problem (3). We let the time horizon be $T_N = tN$ for some fixed $t > 0$, and, for some given $x \in \mathbb{R}_+^2$, we let the starting point $x(N)$ satisfy $\lim_{N\to\infty} x(N)/N = x$. The following result relates the values of the discrete optimal control and the fluid optimal control. For a proof, see Gajrat and Hordijk (2000a).

THEOREM 1.

$$V_{\text{opt}}(x(N)) = N^2 F_{\text{opt}}(x) + o(N^2),$$

where $\lim_{N\to\infty} x(N)/N = x$.

This result leads naturally to a definition of asymptotically fluid optimal policies.

DEFINITION 1. Any policy a for which

$$V_a(x(N)) = N^2 F_{\text{opt}}(x) + o(N^2)$$

is called an asymptotically fluid optimal (a.f.o.) policy.

The problem of finding an a.f.o. policy can be nontrivial, in particular, for the set of parameters (4). For instance, suppose that we translate the optimal fluid control (5) to the following policy in the stochastic network:

$$\begin{aligned} &\text{whenever } \xi_1^n > 0 \text{ and } \xi_2^n > 0: && a_1^n = 0,\ a_2^n = 1,\\ &\text{whenever } \xi_1^n > 0 \text{ and } \xi_2^n = 0: && a_1^n = 1,\ a_2^n = 0,\\ &\text{whenever } \xi_1^n = 0 \text{ and } \xi_2^n = 0: && a_1^n = 0,\ a_2^n = 0. \end{aligned} \tag{7}$$

Then the fluid limit of the trajectories satisfies the equations:

$$\dot{x}^s = \begin{cases} (\lambda, -\mu_2), & \text{if } x_1^s > 0 \text{ and } x_2^s > 0,\\ \left(\lambda - \dfrac{\mu_1\mu_2}{\mu_1 + \mu_2}, 0\right), & \text{if } x_1^s > 0 \text{ and } x_2^s = 0,\\ (0, 0), & \text{if } x_1^s = 0 \text{ and } x_2^s = 0. \end{cases}$$

Clearly, this is not the optimal trajectory (5); thus, the policy defined in (7) is not an a.f.o. policy. However, for this specific queueing model, Gajrat and Hordijk (2000a) considered the following policy. Let $\gamma > 0$ be some positive parameter and let $h(x_1) = [\gamma \ln x_1]$ be a function on the x_1 boundary of the state space ($[x]$ denotes the largest integer not larger than x). Define (see also Figure 1):

$$\begin{aligned} &\text{whenever } \xi_1^n > 0 \text{ and } \xi_2^n > h(\xi_1^n): && a_1^n = 0,\ a_2^n = 1,\\ &\text{whenever } \xi_1^n > 0 \text{ and } \xi_2^n \le h(\xi_1^n): && a_1^n = 1,\ a_2^n = 0,\\ &\text{whenever } \xi_1^n = 0 \text{ and } \xi_2^n = 0: && a_1^n = 0,\ a_2^n = 0. \end{aligned} \tag{8}$$

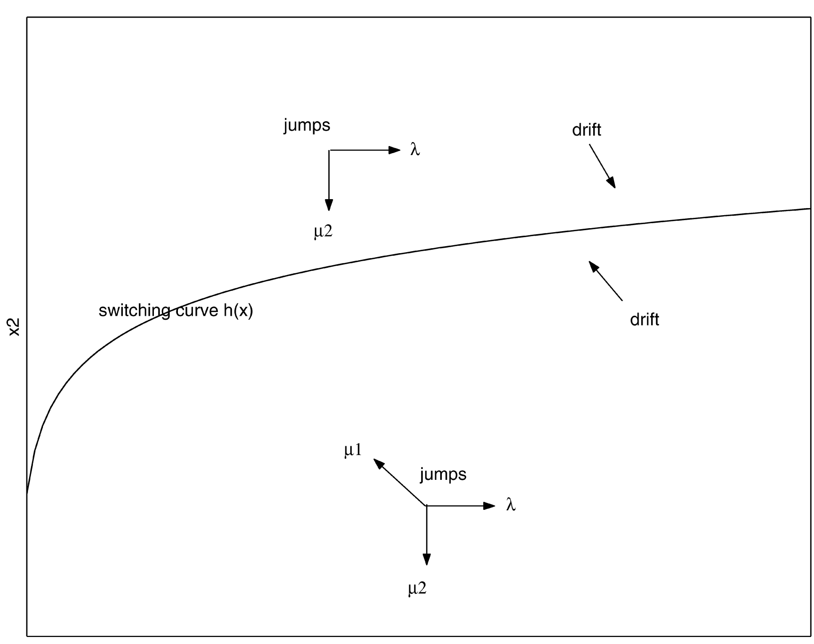


FIG. 1. *The setting of a two-dimensional random walk with switching curve.*

So the first server is idle whenever the number of customers in the second buffer x_2 is greater than $h(x_1)$. The function h is a *switching curve*. For the set of parameters (4), Gajrat and Hordijk (2000a) proved that the policy (8) is asymptotically fluid optimal and that the random walk $\{\xi^n\}$, when starting in $\xi^0 = x(N)$, satisfies

$$\lim_{N\to\infty} \xi^{Ns}/N = x^s_{\text{opt}}.$$

REMARK 2. It is not necessary to choose the switching curve h as a logarithmic function; it can be any smooth (smoothness is convenient only for technical reasons) sublinear function on $\mathbb{R}_{>0}$, meaning that

$$\lim_{x\to\infty} h(x) = \infty, \qquad \lim_{x\to\infty} \frac{h(x)}{x} = 0. \tag{9}$$

Any switching curve $h(x)$ satisfying (9) gives an asymptotically optimal policy. But if we consider the next term in the asymptotic, then the choice of a logarithmic curve gives the smallest asymptotic (see Remark 5).

For different values of $\gamma > 0$, we will have the same first term of the asymptotic for the value function. In the next section, we will consider how the next term depends on γ.

3. Next-order approximation. The main result of the paper is the following.

THEOREM 2. *Let* (4) *hold for the parameters, let* $h(x_1) = [\gamma \ln x_1]$ *be a switching curve and let the control* a *be defined by* (8). *Let the time horizon of the random walk be* $T_N = tN$ *for some fixed* $t > 0$ *and let the initial state be* $x(N) = (x_1(N), x_2(N))$ *with*

$$x_1(N) = [Nx_1], \qquad x_2(N) = h(x_1(N))$$

for some $x_1 > t$. *Then there is a constant*

$$\alpha = \ln \frac{\mu_1(1-\mu_2)}{\mu_2(1-\mu_1)} > 0$$

such that the value function $V_a(x(N))$ *defined in* (2) *satisfies*

$$V_a(x(N)) = N^2 F_{\text{opt}}(x) + N N^{1-\alpha\gamma+o(1)} + c_2 t\gamma N \ln(N) + O(N),$$

where

$$x = (x_1, 0) = \lim_{n\to\infty} x(N)/N.$$

REMARK 3. We use the condition $t < x_1$ just to simplify the proof. It ensures that the random walk $\{\xi^n\}$ never reaches the origin 0.

Hence, we get:

- for $\gamma > 1/\alpha$,

$$\lim_{N\to\infty} \bigl(V_a(x(N)) - N^2 F_{\text{opt}}(x)\bigr)/(N\ln(N)) = c_2 t\gamma,$$

  in words, the next term in the asymptotic after N^2 is $c_2 t\gamma N \ln(N)$;
- for $\gamma < 1/\alpha$,

$$\lim_{N\to\infty} \ln\bigl(V_a(x(N)) - N^2 F_{\text{opt}}(x)\bigr)/\ln(N) = 2 - \alpha\gamma,$$

  the next term is $N^{2-\alpha\gamma} N^{o(1)}$.

So, by decreasing the value of the parameter γ [meaning that the switching curve h lies lower and that the value function $V_a(x(N))$ becomes less] until $1/\alpha$, we see a jump of the asymptotic to very high values.

REMARK 4 (On tracking policies). It is interesting to compare the asymptotic of the policy defined by $h(x)$ to the asymptotic of a policy that is similar to the tracking policy defined in Bäuerle (2000). There, the tracking policies are defined for a class of stochastic networks that differ slightly from our tandem model. They are continuous in time and have an action set, where it is allowed to change service rates. But it is not difficult to give a similar construction of such a policy in the case of our discrete-time tandem model. Let the sequence of initial states $x(N)$ be the same as in Theorem 2 and let $t < x_1$. In this case, the tracking policy corresponds to the tandem network with modified probability of serving in the first buffer $\tilde{\mu}_1 = \mu_1 u_1 = \mu_2$, $u_1 = \mu_2/\mu_1$ and

$$\text{whenever } \xi_2^n > 0: \qquad a_1^n = 1,\ a_2^n = 1,$$
$$\text{whenever } \xi_2^n = 0: \qquad a_1^n = 1,\ a_2^n = 0.$$

So, in the interior part of $\mathbb{Z}_+^2$, ξ^n is a homogeneous random walk with zero vertical drift:

$$E\xi^n = (\lambda - \mu_2, \mu_2 - \mu_2) = (\lambda - \mu_2, 0).$$

It is not too difficult to see that for such a policy the asymptotic of the value function will be

$$V_a(x(N)) = N^2 F_{\text{opt}}(x) + \text{const} \cdot N^{1+1/2+o(1)} \tag{10}$$

for the same sequence of $x(N)$ as in Theorem 2. So the theorem shows that a threshold type of policy gives a better asymptotic than the tracking policies.

We do not give the proof of estimation (10), but the intuition is rather clear. The vertical component ξ_2^n behaves like a random walk with zero drift if $\xi_2^n > 0$ and reflection in 0. So $\xi_2^{tN} \approx N^{1/2}$ (central limit theorem); hence, the contribution of this term to the value function for the time interval $[0, tN]$ will be $\approx N^{1+1/2}$.

4. Proof of Theorem 2. In this section, we assume the following model parameters:

- The stochastic process of the tandem queue is the random walk $\{\xi^0, \xi^1, \dots\} \subset \mathbb{Z}^2_{\geq 0}$.
- The policy a of the tandem queue is defined by (8) in Section 2.2 and is determined by the switching curve $h(t) = [\gamma \ln(t)]$, $t \in \mathbb{R}_{>0}$.
- The initial state of the random walk $x(N) = (x_1(N), x_2(N))$ lies on the switching curve: $x_1(N) = [x_1 N]$ with $x_1 \in \mathbb{R}_{>0}$ and $x_2(N) = h(x_1(N))$.
- The time horizon is $T_N = tN$, where t is called the scaled time horizon satisfying $t < x_1$.
- The scaling factor is $N = 1, 2, \dots$.
- The limit point of the sequence $(x(N)/N)_{N=1}^{\infty}$ is $x = (x_1, 0)$.

In the previous sections, we sub- and superscripted probabilities and expectations in order to denote explicitly their dependence on the initial state and policy: $P^a_{x(N)}$ and $E^a_{x(N)}$. Bearing this in mind, we delete these sub- and superscripts from now on, except in some proofs where we need them. Also, conveniently, we write $x_1 N$ whenever we mean the integer $[x_1 N]$.

4.1. *The value function.* The first thing we do is rewrite the value function $V(x(N))$ as a sum of four terms $V_i(x(N))$, $i = 1, 2, 3, 4$, whose asymptotic behaviors we will analyze subsequently. We need the stochastic variables ν_n, $n = 1, 2, \dots$, defined as the number of times the process visits the x_2-boundary (until time epoch n):

$$\nu_n := \sum_{k=0}^{n-1} \mathbb{1}(\xi_2^k = 0).$$

LEMMA 1.

$$V(x(N)) = V_1(x(N)) + V_2(x(N)) + V_3(x(N)) + V_4(x(N)),$$

where

$$V_1(x(N)) = \sum_{n=0}^{tN} c_1\big(x_1(N) + n(\lambda - \mu_2)\big), \tag{11}$$

$$V_2(x(N)) = c_1 \sum_{n=0}^{tN} \big(x_2(N) - E\xi_2^n\big), \tag{12}$$

$$V_3(x(N)) = c_1 \mu_2 \sum_{n=0}^{tN} E\nu_n, \tag{13}$$

$$V_4(x(N)) = c_2 \sum_{n=0}^{tN} E\xi_2^n. \tag{14}$$

PROOF. The value function is the expected total cost up to time tN. Using the linear property of expectation,

$$V(x(N)) = c_1 \sum_{n=0}^{tN} E\xi_1^n + c_2 \sum_{n=0}^{tN} E\xi_2^n.$$

Observing the process from its initial state, we have

$$E\xi_1^n = x_1(N) + \sum_{k=0}^{n-1} E(\xi_1^{k+1} - \xi_1^k),$$

and when we apply (1):

$$E(\xi_1^{k+1} - \xi_1^k) = \lambda - \mu_1 E a_1(\xi^k).$$

Similarly for the second coordinate:

$$E\xi_2^n = x_2(N) + \sum_{k=0}^{n-1} E(\xi_2^{k+1} - \xi_2^k),$$

$$E(\xi_2^{k+1} - \xi_2^k) = \mu_1 E a_1(\xi^k) - \mu_2 E a_2(\xi^k).$$

Combining these expressions, we get

$$\begin{aligned}\sum_{k=0}^{n-1} E(\xi_1^{k+1} - \xi_1^k) &= \sum_{k=0}^{n-1} \lambda - \sum_{k=0}^{n-1} \mu_1 E a_1(\xi^k)\\ &= n\lambda - \sum_{k=0}^{n-1} E(\xi_2^{k+1} - \xi_2^k) - \sum_{k=0}^{n-1} \mu_2 E a_2(\xi^k)\\ &= n\lambda - \bigl(E\xi_2^n - x_2(N)\bigr) - \sum_{k=0}^{n-1} \mu_2 E a_2(\xi^k).\end{aligned}$$

Thus, for each $n = 1, 2, \dots,$ the first coordinate of the process satisfies

$$E\xi_1^n = x_1(N) + n\lambda + \bigl(x_2(N) - E\xi_2^n\bigr) - \sum_{k=0}^{n-1} \mu_2 E a_2(\xi^k).$$

Now, we apply the fact that in any state of the process the second server either works or is idle:

$$E a_2(\xi^k) + E\mathbb{1}(\xi_2^k = 0) = 1 \qquad \text{for all } k = 0, 1, \dots .$$

So,

$$\begin{aligned}E\xi_1^n &= x_1(N) + n\lambda + \bigl(x_2(N) - E\xi_2^n\bigr) - \sum_{k=0}^{n-1} \mu_2\bigl(1 - E\mathbb{1}(\xi_2^k = 0)\bigr)\\ &= x_1(N) + n(\lambda - \mu_2) + \bigl(x_2(N) - E\xi_2^n\bigr) + \mu_2 E\nu_n.\end{aligned}$$

Multiplying by the cost vector c and adding terms $n = 0, 1, \dots, tN$, we get the four terms (11)–(14). □

Let us try to describe these four terms of the decomposition of the value function. The last term is the expected total cost at the second buffer during the planning horizon; the expected cost at the first buffer is split into three parts corresponding to the first three terms. The first term is related to the value (6) of the optimal control in the fluid model (we come back to this); in fact, it is asymptotically equal to $N^2 F_{\mathrm{opt}}(x)$. Since the optimal fluid trajectory starts in the boundary point $(x_1, 0)$ and moves continuously along the boundary into the direction of 0 $(0, 0)$, this first term takes into account only cost when the random walk drifts along the boundary to 0. Therefore, we need to supply the expected total cost at buffer 1 in the other situations of our discrete stochastic model. Being somewhat disguised, but from the manipulations in the proof of Lemma 1, we infer that the second term deals with the expected cost at buffer 1 when buffer 2 is positive and the third term when buffer 2 is empty and remains empty, that is, no service completion of server 1.

REMARK 5. We give here a heuristic argument as to why we choose a logarithmic function $h(x)$. Because the drift for the random walk is directed toward the graph of $h(x)$, the position of ξ^n will be around $h(x)$, so $E_x^a \xi_2^n \approx h(N)$, and we can expect that $E_x^a \nu_{tN} \approx N \exp(-Ch(N))$ (for some $C > 0$). Hence, $h(x)$ should be chosen in such a way that both terms $E_x^a \nu_n$ and $E_x^a \xi_2^n$ are comparable so $N \exp(-Ch(N)) \approx h(N)$ or $Ch(N) \approx \ln N - \ln h(N)$, but if $h(N)$ is sublinear then $\ln h(N) = o(\ln N)$.

The way to continue is by finding asymptotics of each of the terms (11)–(14). As mentioned above, from the first term we get the fluid value. The fourth term is asymptotically equal to $c_2 t N \gamma \ln(N)$, which we might explain heuristically as follows. The random walk ξ^n starts off at the switching curve with the drift of the process being toward and along the switching curve. This curve is flat for states lying far off 0 while we assume that the process never reaches 0 by taking the time horizon small enough. So we get $\xi_2^n \approx h(x_1 N) = \gamma \ln(x_1) + \gamma \ln(N)$ for $n = 1, 2, \dots, tN$, in which case the second term dominates and leads to the claimed asymptotics. With this asymptotics, the second term (12) disappears, leaving the third term. That is the hard part and we will describe it in more detail in Sections 4.5 and 4.6.

4.2. *Term* 1: *the fluid value.* Notice that

$$\sum_{n=0}^{tN} x_1(N) = \sum_{n=0}^{tN} x_1 N = x_1 t N^2 = N^2 \int_0^t x_1 \, ds$$

and

$$\sum_{n=0}^{tN} n = \int_0^t N[sN]\,ds = N^2 \int_0^t [sN]/N\,ds = N^2\left(\int_0^t s\,ds + O(1/N)\right),$$

when $N \to \infty$. Since the optimal fluid value is $F_{\mathrm{opt}}(x) = \int_0^t c_1(x_1 + s(\lambda - \mu_2))\,ds$, we clearly get

$$V_1(x(N)) = N^2 F_{\mathrm{opt}}(x) + O(N). \tag{15}$$

4.3. *Term* 4: *the switching curve.* The asymptotic of the fourth term (14) follows from the following lemma.

LEMMA 2.

$$\sum_{n=0}^{tN} E\xi_2^n = tN\gamma\ln(N) + O(N).$$

PROOF. Consider the process $\zeta^n = \xi_2^n - h(\xi_1^n)$. If $\zeta^n < 0$, then

$$\begin{aligned} E(\zeta^{n+1} - \zeta^n | \zeta^n) &= E\big(h(\xi_1^n) - h(\xi_1^{n+1}) | \zeta^n\big) \\ &\qquad + E\big(\xi_2^{n+1} - \xi_2^n | a_1(\xi^n) = 1, a_2(\xi^n) = 1\big) \\ &= \mu_1 - \mu_2 + h(\xi_1^n) E\big(\ln\big(1 + (\xi_1^{n+1} - \xi_1^n)/\xi_1^n\big) | \zeta^n\big) \\ &= \mu_1 - \mu_2 + o(1) > 0. \end{aligned}$$

If $\zeta^n > 0$, then

$$E(\zeta^{n+1} - \zeta^n | \zeta^n) = -\mu_2 + o(1) < 0.$$

Hence,

$$P\big(|h(\xi_1^n) - \xi_2^n| > K\big) \le C_1 \exp(-C_2 * K) \tag{16}$$

for some $C_1, C_2 > 0$. Rewrite

$$\sum_{n=0}^{tN} E\xi_2^n = N \int_0^t E\xi_2^{[sN]}\,ds.$$

From (16), we get that, for all $s \in [0, t]$,

$$E\xi_2^{[sN]} = Eh\big(\xi_1^{[sN]}\big) + O(1),$$

where $\xi_1^{[sN]} \in [(x_1 - s)N, (x_1 + s)N]$. Thus, we conclude that $E\xi_2^{[sN]} = h(N) + O(1)$ for all s, or

$$N \int_0^t E\xi_2^{[sN]}\,ds = tN\gamma\ln(N) + O(N). \qquad \square$$

The resulting asymptotic of the value function term reads as

$$V_4(x(N)) = c_2 tN\gamma\ln(N) + O(N). \tag{17}$$

4.4. *Term* 2: *the vanishing part.* This one is easy. The expression (12) of the second term is

$$V_2(x(N)) = c_1 \sum_{n=0}^{tN} x_2(N) - c_1 \sum_{n=0}^{tN} E\xi_2^n.$$

Now, simply use $x_2(N) = h(x_1 N) = \gamma \ln(N) + O(1)$ and apply the asymptotics of the previous section to get

$$V_2(x(N)) = O(N). \tag{18}$$

4.5. *Term* 3: *the* $\{x_2 = 0\}$*-boundary.* The third term (13) of the decomposition of the value function is the most important one. It gives us the "next-order asymptotics." Here is what it is all about.

LEMMA 3.

$$\sum_{n=0}^{tN} E\nu_n = N^{2-\alpha\gamma+o(1)},$$

with

$$\alpha = \ln \frac{\mu_1(1-\mu_2)}{\mu_2(1-\mu_1)}.$$

PROOF. We start off doing some calculus manipulations:

$$\begin{aligned}
\sum_{n=0}^{tN} E\nu_n &= \sum_{n=0}^{tN} \sum_{k=0}^{n-1} P(\xi_2^k = 0) \\
&= \sum_{k=0}^{tN-1} \sum_{n=k+1}^{tN} P(\xi_2^k = 0) \\
&= \sum_{k=0}^{tN-1} (tN-k) P(\xi_2^k = 0) \\
&= N \int_0^{t-(1/N)} (tN - [sN]) P\big(\xi_2^{[sN]} = 0\big)\, ds \\
&= N^2 \int_0^{t-(1/N)} (t - [sN]/N) P\big(\xi_2^{[sN]} = 0\big)\, ds \\
&= N^2 \int_0^{t} (t-s) P\big(\xi_2^{[sN]} = 0\big)\, ds + O(N).
\end{aligned} \tag{19}$$

All we have to do now is to find asymptotics for $P(\xi_2^{[sN]} = 0)$ and that will be the subject of the next section. Here we just state the results from Lemmas 4 and 5 (forthcoming):

- For $s > t - \varepsilon$ (with $\varepsilon > 0$ arbitrary but small),

$$P(\xi_2^{[sN]} = 0) \geq N^{-\alpha\gamma + o(1)}.$$

- For $s > rh(x_1 N)/N$ [with $r > 0$ arbitrary but less than $tN/h(x_1 N)$],

$$P(\xi_2^{[sN]} = 0) \leq N^{-\alpha\gamma + o(1)}.$$

Thus, we can bound the integral in (19):

$$\begin{aligned}\int_0^t (t-s) P(\xi_2^{[sN]} = 0)\, ds &\geq \int_{t-\varepsilon}^t (t-s) P(\xi_2^{[sN]} = 0)\, ds \\ &\geq N^{-\alpha\gamma + o(1)}\end{aligned}$$

and

$$\begin{aligned}\int_0^t (t-s) P(\xi_2^{[sN]} = 0)\, ds &\leq \int_{rh(x_1 N)/N}^t (t-s) P(\xi_2^{[sN]} = 0)\, ds \\ &\quad + \int_0^{rh(x_1 N)/N} (t-s)\, ds \\ &\leq N^{-\alpha\gamma + o(1)} + O(\ln(N)/N).\end{aligned}$$

So,

$$N^2 \int_0^t (t-s) P(\xi_2^{[sN]} = 0)\, ds = N^{2-\alpha\gamma + o(1)}. \qquad \square$$

Thus, the resulting asymptotic of the third term of the value function is

$$V_3(x(N)) = N^{2-\alpha\gamma + o(1)} + O(N). \tag{20}$$

The asymptotics (15), (17), (18) and (20) of the separate terms of the value function add up to the result claimed in Theorem 2.

4.6. *Asymptotics for boundary probabilities.* We give an asymptotic of $P(\xi_2^{[sN]} = 0)$ for large N based on large deviations. For notational convenience, we write $(\xi^{\sigma N})_{\sigma \geq 0}$ to mean the random walk $\{\xi^n\}_{n=0,1,\ldots}$ by assuming that σ takes values in the scaled integers $\{0, 1/N, 2/N, \ldots\}$. Let us first sketch the idea of the proof (see also Figure 2). Consider trajectories that reach the $\{x_2 = 0\}$-boundary at time sN. During the first part of the trajectory, the random walk wanders around the switching curve in a neighborhood of the starting point. The switching curve is quite flat there; thus, the second coordinate $\xi_2^{\sigma N}$ remains approximately equal to the original level $h(x_1 N)$. At a particular time $\tilde{t}N$, the walk starts to move downward to the boundary, and therefore the statistical law of the vertical movement (i.e., the second coordinate) of the random walk is most relevant. The probabilities that the random walk goes up or down are exactly $\tilde{\mu}_1 := \mu_1(1 - \mu_2)$ and

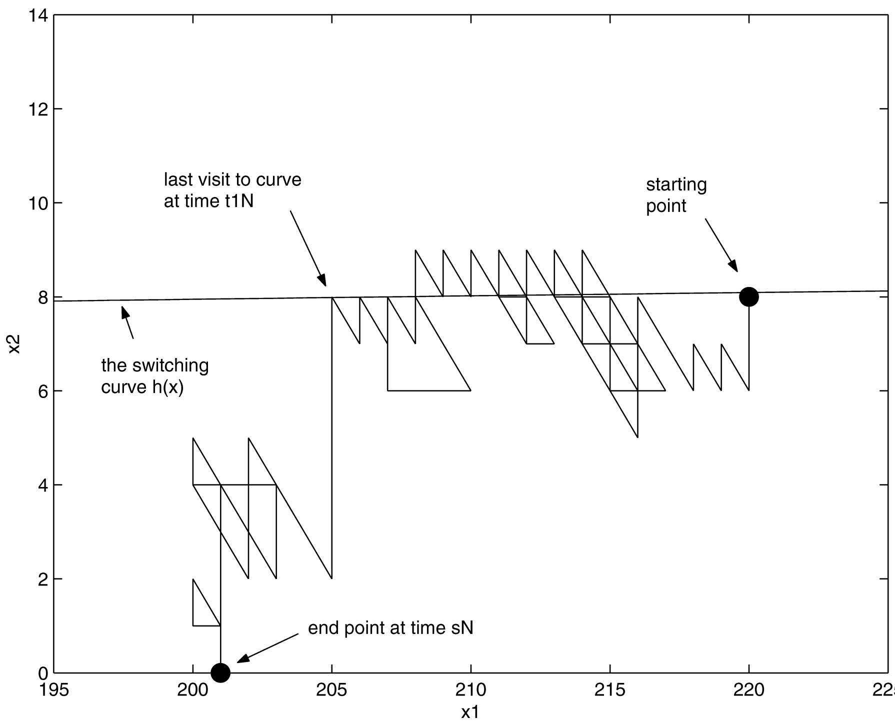


FIG. 2. *The random walk until hitting the boundary at time sN.*

$\tilde{\mu}_2 := \mu_2(1-\mu_1)$, respectively. Notice that $\tilde{\mu}_1 > \tilde{\mu}_2$ because $\mu_1 > \mu_2$, and that this again says that the random walk tends to jump toward the switching curve whenever it resorts below it. The mean time for the random walk to cross from the $\{x_2 = 0\}$-boundary to the switching curve is $h(x_1 N)/(\tilde{\mu}_1 - \tilde{\mu}_2)$ (just following the drift). Applying large-deviations theory for random walks, it turns out that exactly the same time is most likely to occur when the random walk jumps in the opposite direction (from the curve to the boundary). That is, the last visit to the switching curve is most likely at time $sN - h(x_1 N)/(\tilde{\mu}_1 - \tilde{\mu}_2)$. In the proofs of the asymptotics, we will see this number occurring. We will discuss the large-deviations material in Section 4.8.

Partition $\mathbb{Z}_+^2$ into two regions: $A^+ := \{(x_1, x_2) \in \mathbb{Z}_+^2 : x_2 \geq h(x_1)\}$, the set of all points on and above the switching curve, and $A^- := \{(x_1, x_2) \in \mathbb{Z}_+^2 : x_2 < h(x_1)\}$, the set of points below the switching curve. We make the observation that in order for the random walk $(\xi^{\sigma N})_{\sigma \geq 0}$ to reach the $\{x_2 = 0\}$-boundary at time sN, it must have walked through the set A^- during a time interval $(\tau N, sN]$. Here, τN is the (random) time epoch of the last visit to the set A^+. Suppose $\tau = \tilde{t}$ and let $\tilde{x}(N) \in A^+$ be the last state visited; hence, $\tilde{x}(N) = (\tilde{x}_1 N, h(\tilde{x}_1 N))$ lies on the switching curve. Since we assume the scaled horizon t to be small relative to x_1 of the starting state $(x_1 N, h(x_1 N))$, the second coordinate of these points satisfies

$h(\tilde{x}_1 N) \approx h(x_1 N)$. This means that the remaining time $(s - \tilde{t})N$ that is left for the random walk to go all the way down to the boundary takes at least $h(x_1 N)$ time units. In other words, the exit time should satisfy $\tau N \leq sN - h(x_1 N)$. Also, it means that $P(\xi_2^{sN} = 0) = 0$ for all $sN \leq h(x_1 N)$. Now, the idea is to find the asymptotic of $P(\xi_2^{sN} = 0)$ by lower and upper bounding it.

Let us restate some notation:

- The random walk is $\{\xi^{\sigma N} : \sigma = 0, 1/N, 2/N, \ldots\}$.
- The initial point $x(N) = (x_1 N, h(x_1 N))$ lies on the switching curve.
- The planning horizon is tN with $t < x_1$.
- τN is the last time that the random walk visits the area on or above the switching curve. We call the random variable τ the scaled exit time.
- Realizations of τ are denoted by $\tilde{t}$. Realizations of the last state visited on or above the switching curve are denoted by $\tilde{x}(N)$, from where the random walk starts moving downward to approach the boundary.
- $\tilde{\mu}_1 = \mu_1(1 - \mu_2)$ and $\tilde{\mu}_2 = \mu_2(1 - \mu_1)$. These are the probabilities that in the tandem queue server 1 serves and server 2 does not, and, respectively, server 2 serves and server 1 does not.
- The "speed" at which the random walk moves upward is $\mu = 1/(\tilde{\mu}_1 - \tilde{\mu}_2)$. Notice that $\mu > 1$.
- $\alpha = \ln(\tilde{\mu}_1/\tilde{\mu}_2)$.

The two bounding statements are as follows.

LEMMA 4. *Let $\varepsilon > 0$ small. Then, for $s > t - \varepsilon$,*

$$P(\xi_2^{sN} = 0) \geq N^{-\alpha\gamma + o(1)}.$$

LEMMA 5. *Let $r > \mu$, $r < tN/h(x_1 N)$. Then, for $s > rh(x_1 N)/N$,*

$$P(\xi_2^{sN} = 0) \leq N^{-\alpha\gamma + o(1)}.$$

PROOF OF LEMMA 4. Let $s \in (t - \varepsilon, t)$ and consider the time $\tilde{t} = s - h(x_1 N)\mu/N$. We obtain a lower bound of the probability by restricting the scaled exit time τ to realize only this specific $\tilde{t}$:

$$P(\xi_2^{sN} = 0) \geq P(\xi_2^{sN} = 0;\ \tau = \tilde{t}).$$

We partition the right-hand side into all possible exit states on the curve. At exit time $\tilde{t}N$, the random walk is at some (random) state $\xi^{\tilde{t}N}$ on the switching curve. Since the random walk makes jumps of size 1, we know that the first coordinate $\xi_1^{\tilde{t}N} \in [x_1 - \tilde{t}, x_1 + \tilde{t}]N$ with probability 1. We denote by $\tilde{x}(N) = (\tilde{x}_1 N, h(\tilde{x}_1 N))$ these possible states on the switching curve. Hence,

$$P(\xi_2^{sN} = 0;\ \tau = \tilde{t}) = \sum_{\tilde{x}_1 = x_1 - \tilde{t}}^{x_1 + \tilde{t}} P\big(\xi_2^{sN} = 0;\ \tau = \tilde{t};\ \xi^{\tilde{t}N} = \tilde{x}(N)\big).$$

The summands on the right-hand side may be rewritten using the Markov property of the random walk:

$$P_{x(N)}\big(\xi_2^{sN}=0;\ \tau=\tilde t;\ \xi^{\tilde t N}=\tilde x(N)\big)$$
$$=P_{x(N)}\big(\xi^{\tilde t N}=\tilde x(N)\big)\,P_{\tilde x(N)}\big(\xi_2^{(s-\tilde t)N}=0;\ \tau=0\big).$$

The following asymptotic will be proved in Lemma 6(ii). For any $\tilde x(N)$ with $\tilde x_1\in[x_1-\tilde t, x_1+\tilde t\,]$,

$$P_{\tilde x(N)}\big(\xi_2^{(s-\tilde t)N}=0;\ \tau=0\big)\ge N^{-\alpha\gamma+o(1)}$$

[notice that $(s-\tilde t\,)N=\mu h(x_1N)$]. Hence,

$$P_{x(N)}(\xi_2^{sN}=0)\ge N^{-\alpha\gamma+o(1)}\sum_{\tilde x_1=x_1-\tilde t}^{x_1+\tilde t}P_{x(N)}\big(\xi^{\tilde t N}=\tilde x(N)\big)$$
$$=N^{-\alpha\gamma+o(1)}P_{x(N)}\big(\xi_2^{\tilde t N}=[h(\xi_1^{\tilde t N})]\big).$$

The following asymptotic will be proved in Lemma 7:

$$P_{x(N)}\big(\xi_2^{\tilde t N}=[h(\xi_1^{\tilde t N})]\big)=O(1).$$

Putting it all together, we get

$$P_{x(N)}(\xi_2^{sN}=0)\ge N^{-\alpha\gamma+o(1)}. \qquad \square$$

PROOF OF LEMMA 5. Let $r>\mu$ and $s>rh(x_1N)/N$. Then, by decomposition to the values of the exit time,

$$P(\xi_2^{sN}=0)=P\big(\xi_2^{sN}=0;\ \tau N\le sN-rh(x_1N)\big)+P\big(\xi_2^{sN}=0\,;\ \tau N>sN-rh(x_1N)\big). \tag{21}$$

Because the random walk wanders below the switching curve after the exit time, the first term is bounded by

$$P\big(\xi_2^{sN}=0;\ \tau N\le sN-rh(x_1N)\big)$$
$$\le P\big(\xi_2^{\sigma N}<h(\xi_1^{\sigma N})\text{ for all }\sigma N>sN-rh(x_1N)\big).$$

Using the same argument as in the proof of Lemma 2, we have

$$P\big(\xi_2^{\sigma N}<h(\xi_1^{\sigma N})\text{ for all }\sigma N>sN-rh(x_1N)\big)$$
$$\le C_1\exp\big(-C_2rh(x_1N)\big)=C_1N^{-r\gamma C_2}$$

for some constants $C_1,C_2>0$. So, choosing r sufficiently large, we can make this term arbitrarily small. Notice that this is possible because we require $r<tN/h(x_1N)$, which diverges to ∞ when $N\to\infty$.

For the second term on the right-hand side of (21), we have, by partitioning and the Markov property,

$$P_{x(N)}\big(\xi_2^{sN}=0\,;\ \tau N>sN-rh(x_1N)\big)$$
$$=N\int_{s-rh(x_1N)/N}^{s}\sum_{\tilde{x}_1=x_1-\tilde{t}}^{x_1+\tilde{t}}P_{\tilde{x}(N)}\big(\xi_2^{(s-\tilde{t})N}=0;\ \tau=0\big)$$
$$\times P_{x(N)}\big(\xi^{\tilde{t}N}=\tilde{x}(N)\big)\,d\tilde{t}.$$

The following asymptotic will be proved in Lemma 6(i). For any $\tilde{t}>rh(x_1N)/N$ and $\tilde{x}(N)$ with $\tilde{x}_1\in[x_1-s,x_1+s]$,

$$P_{\tilde{x}(N)}\big(\xi_2^{(s-\tilde{t})N}=0;\ \tau=0\big)\le N^{-\alpha\gamma+o(1)}$$

[notice that $(s-\tilde{t})N<rh(x_1N)$]. Hence,

$$P_{x(N)}\big(\xi_2^{sN}=0;\ \tau N>sN-rh(x_1N)\big)$$
$$\le NN^{-\alpha\gamma+o(1)}\int_{s-rh(x_1N)/N}^{s}\sum_{\tilde{x}_1=x_1-\tilde{t}}^{x_1+\tilde{t}}P_{x(N)}\big(\xi^{\tilde{t}N}=\tilde{x}(N)\big)\,d\tilde{t}$$
$$=NN^{-\alpha\gamma+o(1)}\int_{s-rh(x_1N)/N}^{s}P_{x(N)}\big(\xi_1^{\tilde{t}N}=h(\xi_1^{\tilde{t}N})\big)\,d\tilde{t}$$
$$\le rh(x_1N)N^{-\alpha\gamma+o(1)}.$$

Putting it all together and noticing that $\ln(N)=N^{o(1)}$, we get

$$P(\xi_2^{sN}=0)\le N^{-\alpha\gamma+o(1)}.\qquad\square$$

4.7. *More asymptotics to prove.* In this section, we prove the results that were applied in the proofs of Lemmas 4 and 5. The assumptions and notation are the same as listed in Section 4.6. Recall that $\tilde{x}(N)=(\tilde{x}_1N,h(\tilde{x}_1N))$ is a state on the switching curve from where the random walk starts moving downward to approach the boundary, indicated by $\tau=0$. Because we assumed the time horizon tN to be so small relative to the starting point $x(N)$ that the $\{x_1=0\}$-boundary will not be reached, we know $t<x_1$, and thus $\tilde{x}_1\in[x_1-t,x_1+t]$.

LEMMA 6. (i) *Let $r>\mu$. For $\sigma N<rh(x_1N)$,*

$$P_{\tilde{x}(N)}(\xi_2^{\sigma N}=0;\ \tau=0)\le N^{-\alpha\gamma+o(1)}.$$

(ii) *For $\sigma N=\mu h(x_1N)$,*

$$P_{\tilde{x}(N)}(\xi_2^{\sigma N}=0;\ \tau=0)\ge N^{-\alpha\gamma+o(1)}.$$

PROOF. We first show the asymptotics when $\sigma N = \rho h(\tilde{x}_1 N)$ for arbitrary $\rho \in (0, r)$.

Consider a policy $\tilde{a}$ with control variables $\tilde{a}_i^n = 1$ for all states. That is, the switching curve has been removed: both servers service always. However, we consider events where the state of the random walk stays below the curve. Thus,

$$P^{a}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \tau = 0) = P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \tau = 0).$$

Let $\varepsilon > 0$. We will apply large-deviations asymptotics for events involving $|\xi_2^n - h(\tilde{x}_1 N)| < \varepsilon n$, $n = 1, 2, \dots$. Figure 3 shows that

$$\begin{aligned} &P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \tau = 0) \\ &\qquad = P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \xi^n \in \mathrm{A} \cup \mathrm{B} \cup \mathrm{C}, n = 1, 2, \dots, \sigma N), \\ &P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) - \varepsilon n, n = 1, 2, \dots, \sigma N) \\ &\qquad = P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \xi^n \in \mathrm{A}, n = 1, 2, \dots, \sigma N), \\ &P^{\tilde{a}}_{\tilde{x}(N)}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) + \varepsilon n, n = 1, 2, \dots, \sigma N\big) \\ &\qquad = P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \xi^n \in \mathrm{A} \cup \mathrm{B} \cup \mathrm{C} \cup \mathrm{D} \cup \mathrm{E}, n = 1, 2, \dots, \sigma N). \end{aligned}$$

A geometrical argument (see Figure 3) shows the inequalities

$$P^{\tilde{a}}_{\tilde{x}(N)}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) - \varepsilon n, n = 1, 2, \dots, \sigma N\big)$$

$$\le P^{\tilde{a}}_{\tilde{x}(N)}(\xi_2^{\sigma N} = 0, \tau = 0) \tag{22}$$

$$\le P^{\tilde{a}}_{\tilde{x}(N)}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) + \varepsilon n, n = 1, 2, \dots, \sigma N\big). \tag{23}$$

Now, we claim that the upper bound (23) is upper bounded asymptotically:

$$P^{\tilde{a}}_{\tilde{x}(N)}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) + \varepsilon n, n = 1, 2, \dots, \sigma N\big) \le N^{-\alpha\gamma + o(1)} \tag{24}$$

for any $\sigma N = \rho h(\tilde{x}_1 N)$. And the lower bound (22) is asymptotically equal to

$$P^{\tilde{a}}_{\tilde{x}(N)}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) - \varepsilon n, n = 1, 2, \dots, \sigma N\big) = N^{-\alpha\gamma + o(1)}, \tag{25}$$

when $\sigma N = \mu h(\tilde{x}_1 N)$. We will prove these two claims in Section 4.8.

Finally, the question is whether it is necessary to replace $h(\tilde{x}_1 N)$ by $h(x_1 N)$. The answer is no, since $h(\tilde{x}_1 N) - h(x_1 N) = o(\ln(N))$. □

LEMMA 7. *For any $\tilde{t} < x_1$,*

$$P\big(\xi_2^{\tilde{t}N} = h(\xi_1^{\tilde{t}N})\big) = O(1).$$

PROOF. Using (16), we can find $K > 0$ such that

$$P\big(|\xi_2^{\sigma N} - h(\xi_1^{\sigma N})| < K\big) \ge 1/2$$

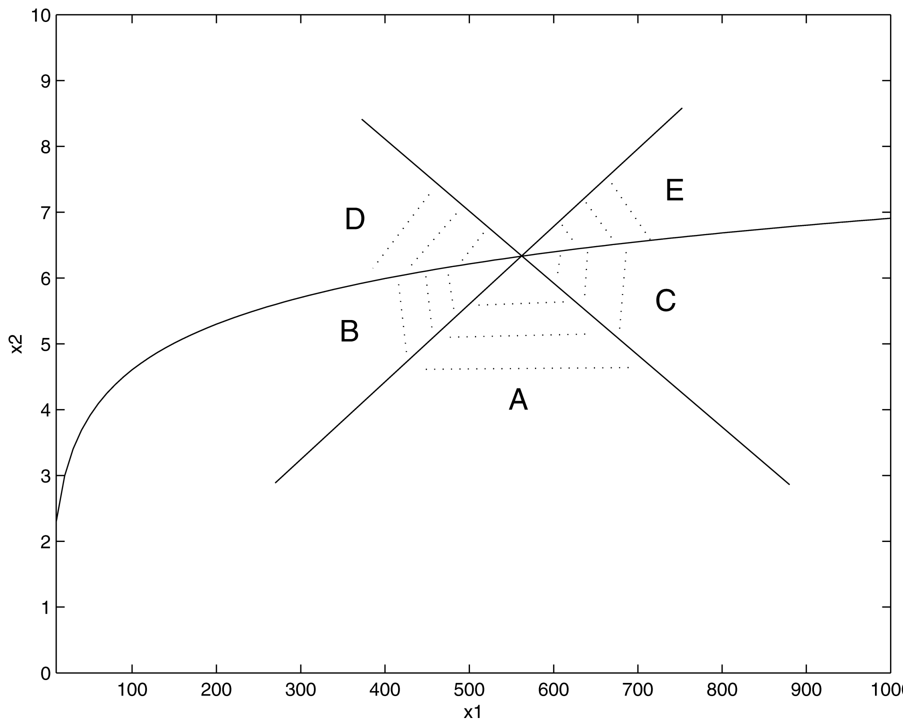


FIG. 3. *The geometric argument.*

for any $\sigma < x_1$. We choose a particular σ to get

$$\begin{aligned} P\big(\xi_2^{\tilde{t}N} = h(\xi_1^{\tilde{t}N})\big) &\geq P\big(|\xi_2^{\tilde{t}N-K} - h(\xi_1^{\tilde{t}N-K})| < K\big) \\ &\qquad \times P\big(\xi_2^{\tilde{t}N} = h(\xi_1^{\tilde{t}N}) \big| |\xi_2^{\tilde{t}N-K} - h(\xi_1^{\tilde{t}N-K})| < K\big) \\ &\geq 1/2\rho^K \end{aligned}$$

for some $\rho < 1$. $\square$

4.8. *Large-deviations leftovers.* Finally, we show the two remaining (in)equalities (24) and (25). Some of the notation is as before:

- probabilities $\tilde{\mu}_1$ and $\tilde{\mu}_2$ with $1 > \tilde{\mu}_1 > \tilde{\mu}_2 > 0$;
- an arbitrary number $\tilde{x}_1 > 0$;
- $\mu = 1/(\tilde{\mu}_1 - \tilde{\mu}_2)$.

Consider a one-dimensional random walk $Y := \{Y^n : n = 0, 1, \dots\}$ on $\mathbb{Z}_{\geq 0}$ with the following jumps and jump probabilities:

(i) When the current state $y \geq 1$:

jump 0	with probability $1 - \tilde{\mu}_1 - \tilde{\mu}_2$,
jump 1	with probability $\tilde{\mu}_1$,
jump -1	with probability $\tilde{\mu}_2$.

(ii) When the current state $y = 0$:

jump 0	with probability $1 - \tilde{\mu_1}$,
jump 1	with probability $\tilde{\mu_1}$.

One may view this random walk as moving along a vertical in the plane. The height at time n is given by Y^n. The random walk has drift $\tilde{\mu}_1 - \tilde{\mu}_2 = 1/\mu$ (in the positive states). Such a process satisfies the large-deviations principle; see, for example, Section 5.1 in Dembo and Zeitouni (1996) or Section 7.2 in Shwartz and Weiss (1995). The jump probabilities are homogeneous, except for the single boundary {0}. The boundary can be taken care of by a reflection map [cf. Section 11.4 in Shwartz and Weiss (1995)]. We apply the large-deviations asymptotic for a special event and by scaling with a special sequence. The scaling sequence is $\{a_N : N = 1, 2, \dots\}$ with $a_N = \ln(\tilde{x}_1 N)$. The event is

$$\{Y^{\rho a_N} < \delta a_N,\ 0 \leq Y^n < y a_N - \varepsilon n,\ n = 1, 2, \dots, \rho a_N\},$$

where $\rho, \delta, \varepsilon > 0$ and $y a_N$ is the initial state of the random walk. In words (see Figure 4): starting from $y a_N$, the random walk stays below the line l (through $y a_N$, with slope $-\varepsilon$) and ends after ρa_N time units at (or close by) the 0-boundary. Since the most likely behavior of the random walk is upward ($\tilde{\mu}_1 > \tilde{\mu}_2$), the probability of this event satisfies a large-deviations asymptotic:

$$\lim_{\delta \downarrow 0} \lim_{N \to \infty} \frac{1}{a_N} \log P_{y a_N}(Y^{\rho a_N} < \delta a_N,\ 0 \leq Y^n < y a_N - \varepsilon n, n = 1, 2, \dots, \rho a_N)$$
$$=: -J(\rho, y, \varepsilon).$$

To determine the rate function $J(\rho, y, \varepsilon)$, we apply the sample path large deviations as treated in Chapter 11 of Shwartz and Weiss (1995). A path is an absolute continuous function $f : [0, \rho] \to \mathbb{R}$. The considered event involves paths $f \in U$ such that

$$f(0) = y, \qquad f(\rho) = 0, \qquad 0 \leq f(t) < y - \varepsilon t \qquad \text{for all } 0 < t < \rho.$$

Then

$$J(\rho, y, \varepsilon) = \inf_{f \in U} \int_0^\rho I\big(f(t), f'(t)\big)\, dt,$$

where $I(\cdot, \cdot)$ is the local rate function. When $\varepsilon < y/\rho$, this variational program is solved for f being a straight line with slope $-y/\rho$ ($\varepsilon < y/\rho$ means that f lies below the line with slope $-\varepsilon$). The rate function $J(\rho, y, \varepsilon)$ is convex unimodal

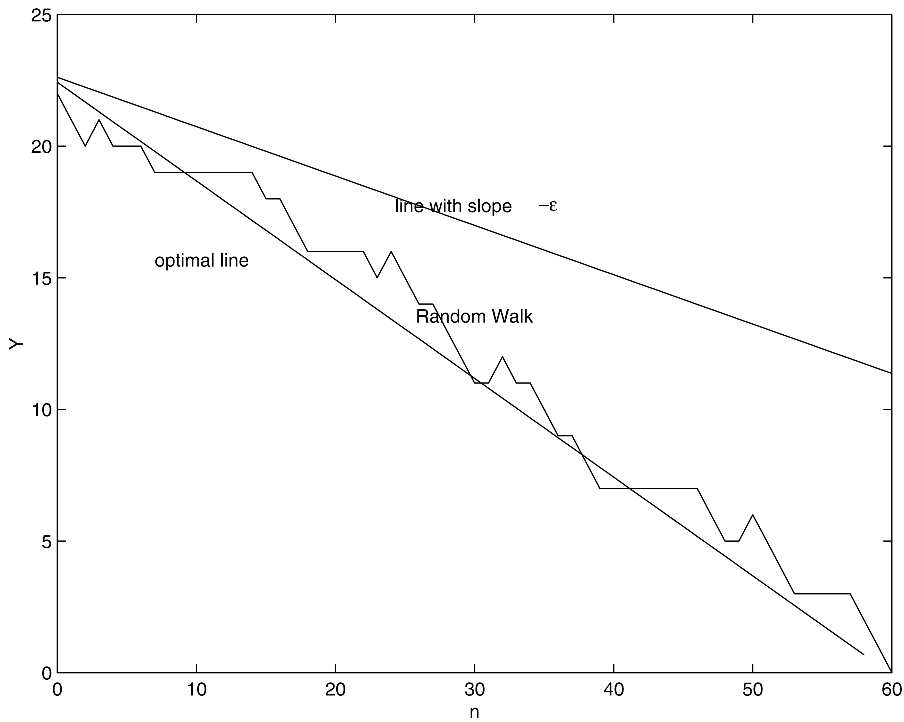


FIG. 4. *Random walk Y^n staying below line l.*

as a function of ρ. The unique minimum is attained at $\rho = \mu y$ (the slope of the optimal f is $\tilde{\mu}_1 - \tilde{\mu}_2$). The optimal rate equals

$$J(\mu, y, \varepsilon) = y \ln \frac{\tilde{\mu}_1}{\tilde{\mu}_2}.$$

Summarizing, we have [with $\alpha := \ln(\tilde{\mu}_1/\tilde{\mu}_2)$]

$$\begin{aligned} &P_{ya_N}(Y^{\rho a_N} = 0,\ 0 \le Y^n < ya_N - \varepsilon n, n = 1, 2, \ldots, \rho a_N) \\ &\qquad = \exp\bigl(-a_N J(\rho, y, \varepsilon) + o(a_N)\bigr) \\ &\qquad \le \exp\bigl(-\alpha y a_N + o(a_N)\bigr), \end{aligned} \tag{26}$$

with equality for $\rho = \mu y$.

The same conclusion can be drawn for the event

$$\bigl\{Y^{\rho a_N} < \delta a_N,\ 0 \le Y^n < ya_N + \varepsilon n, n = 1, 2, \ldots, \rho a_N\bigr\},$$

where $\delta \downarrow 0$. Hence,

$$\begin{aligned} &P_{ya_N}(Y^{\rho a_N} = 0,\ 0 \le Y^n < ya_N + \varepsilon n, n = 1, 2, \ldots, \rho a_N) \\ &\qquad \le \exp\bigl(-\alpha y a_N + o(a_N)\bigr). \end{aligned} \tag{27}$$

Now, let us relate these asymptotics to our problem and show the inequalities (24) and (25). The random walk Y^n stands for the second coordinate ξ_2^n of the tandem queue process when policy $\tilde{a}$ is applied; that is, the servers always serve. Then we set $h(\tilde{x}_1 N) = \gamma a_N$ and $\sigma N = \rho a_N$ to get

$$\begin{aligned} P_{\tilde{x}(N)}^{\tilde{a}}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) - \varepsilon n, n = 1, 2, \dots, \sigma N\big) \\ = P_{\gamma a_N}(Y^{\rho a_N} = 0,\ Y^n < \gamma a_N - \varepsilon n, n = 1, 2, \dots, \rho a_N), \\ P_{\tilde{x}(N)}^{\tilde{a}}\big(\xi_2^{\sigma N} = 0, \xi_2^n < h(\tilde{x}_1 N) + \varepsilon n, n = 1, 2, \dots, \sigma N\big) \\ = P_{\gamma a_N}(Y^{\rho a_N} = 0\,,\ Y^n < \gamma a_N + \varepsilon n, n = 1, 2, \dots, \rho a_N). \end{aligned}$$

Finally, the asymptotics (27) and (26) yield the inequalities (24) and (25), respectively.

5. Optimal switching curve. In this section, we give some computational results (using the value iteration algorithm) for computing an optimal policy of the problem $\min \sum_{t=0}^{\tau} \xi^t c$, where $\tau = \min\{t : \xi^t = (0,0)\}$. This problem is slightly different from the problem of the previous sections, but the same type of asymptotic result can be proved for it. The optimal switching curve C is depicted in Figure 5. The parameters are $\lambda = 0.1$, $\mu_1 = 0.22$, $\mu_2 = 0.2$, $c_1 = 1$, $c_2 = 2$. If we denote by l_n the x_1-coordinate where C hits the level n, then according to our asymptotically best policy with logarithmic switching curve, we would have

$$\lim_{n\to\infty} l_n / l_{n-1} = \exp(\alpha) = \mu_1(1-\mu_2)/\big(\mu_2(1-\mu_1)\big) = 1.12821.$$

The computation gives the following numbers:

$$\begin{aligned} \{l_n\}_{n=5}^{22} = \{&7, 13, 20, 29, 40, 53, 69, 88, 109, 135, \\ &164, 197, 236, 279, 329, 385, 450, 525\} \end{aligned}$$

and

$$\begin{aligned} \{l_n / l_{n-1}\}_{n=6}^{22} = \{&1.8571, 1.5385, 1.45, 1.3793, 1.325, 1.3019, \\ &1.2754, 1.2386, 1.2385, 1.2148, 1.2012, \\ &1.198, 1.1822, 1.1792, 1.1702, 1.1688, 1.1667\}. \end{aligned}$$

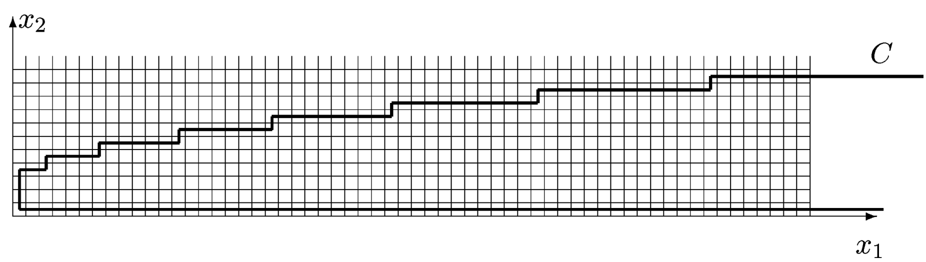


FIG. 5. *The optimal switching curve.*

This seems to indicate that the asymptotic of the switching curve is close to a logarithmic function.

Acknowledgment. The authors thank an anonymous referee for valuable remarks that have improved the paper.

REFERENCES

ATKINS, D. and CHEN, H. (1995). Performance evaluation of scheduling control of queueing network: Fluid model heuristics. *Queueing Systems* **21** 391–413.

AVRAM, F., BERTSIMAS, D. and RICARD, M. (1995). Fluid models of sequencing problems in open queueing networks: An optimal control approach. In *Stochastic Networks* (F. Kelly and R. Williams, eds.) **71** 199–234. Springer, New York.

BÄUERLE, N. (2000). Asymptotic optimality of tracking-policies in stochastic networks. *Ann. Appl. Probab.* **10** 1065–1083.

BÄUERLE, N. and RIEDER, U. (2000). Optimal control of single-server fluid networks. *Queueing Systems* **35** 185–200.

BELL, S. L. and WILLIAMS, R. J. (2001). Dynamic scheduling of a system with two parallel servers in heavy traffic with complete resource pooling: Asymptotic optimality of a continuous review threshold policy. *Ann. Appl. Probab.* **11** 608–649.

BLINOSKII, V. M. and DOBRUSHIN, R. L. (1994). Process level large deviations for a class of piecewise homogeneous random walks. In *The Dynkin Festschrift*: *Markov Processes and Their Applications* (M. I. Friedlin, ed.) 1–59. Birkhäuser, Boston.

DEMBO, A. and ZEITOUNI, O. (1996). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York.

GAJRAT, A., HORDIJK, A., MALYSHEV, V. and SPIEKSMA, F. (1997). Fluid approximations of Markov decision chains. *Markov Process. Related Fields* **3** 129–150.

GAJRAT, A. and HORDIJK, A. (1999). Optimal service control of multiclass fluid tandem networks. Available at www.math.leidenuniv.nl/~gajrat/publications/gh2/tandem2.ps.

GAJRAT, A. and HORDIJK, A. (2000a). Fluid approximation of a controlled multiclass tandem network. *Queueing Systems* **35** 349–380.

GAJRAT, A. and HORDIJK, A. (2000b). Synthesis of the optimal control problem for fluid networks. Available at www.math.leidenuniv.nl/~gajrat/publications/gh3/general12.ps.

HARRISON, J. M. (1996). The BIGSTEP approach to flow management in stochastic processing networks. In *Stochastic Networks*: *Theory and Applications* (F. Kelly, S. Zachart and I. Ziedens, eds.) 57–90. Oxford Univ. Press.

HARRISON, J. M. (1998). Heavy traffic analysis of a system with parallel servers: Asymptotic optimality of discrete-review policies. *Ann. Appl. Probab.* **8** 822–848.

IGNATYUK, I. A. (1998). Large deviations for a random walk in dynamical random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **34** 601–639.

IGNATYUK, I. A., MALYSHEV, V. and SCHERBAKOV, V. V. (1994). Boundary effects in large deviation problems. *Russian Math. Surveys* **49** 41–99.

KELLY, F. and LAWS, N. (1993). Dynamic routing in open queueing networks: Brownian models, cut constraints and resource pooling. *Queueing Systems* **35** 349–380.

MAGLARAS, C. (1999). Dynamic scheduling in multiclass queueing networks: Stability under discrete-review policies. *Queueing Systems* **31** 171–206.

MAGLARAS, C. (2000). Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid scale asymptotic optimality. *Ann. Appl. Probab.* **10** 897–929.

MAGLARAS, C. (2002). Continuous-review tracking policies for dynamic control of stochastic networks. *Queueing Systems*. To appear.

PULLAN, M. C. (1995). Forms of optimal solutions for separated continuous linear programs. *SIAM J. Control Optim.* **33** 1952–1977.

PULLAN, M. C. (1996). A duality theory for separated continuous linear programs. *SIAM J. Control Optim.* **34** 931–965.

SHWARTZ, A. and WEISS, A. (1995). *Large Deviations for Performance Analysis*: *Queues, Communication and Computing*. Chapman and Hall, New York.

WEISS, G. (1995). On optimal draining of fluid re-entrant lines. In *Stochastic Networks* (F. Kelly and R. Williams, eds.) **71** 93–105. Springer, New York.

WEISS, G. (1999). Scheduling and control of manufacturing systems—a fluid approach. In *Proc. 37th Allerton Conference* 557–586.

A. GAJRAT
DEPARTMENT OF ECONOMETRICS
VRIJE UNIVERSTEIT AMSTERDAM
DE BOELELAAN 1105
1081 HV AMSTERDAM
THE NETHERLANDS

A. HORDIJK
MATHEMATICAL INSTITUTE
LEIDEN UNIVERSITY
P.O. BOX 9512
2300 RA LEIDEN
THE NETHERLANDS
E-MAIL: hordijk@math.leidenuniv.nl

A. RIDDER
DEPARTMENT OF ECONOMETRICS
VRIJE UNIVERSTEIT AMSTERDAM
DE BOELELAAN 1105
1081 HV AMSTERDAM
THE NETHERLANDS
E-MAIL: aridder@feweb.vu.nl