

## ASYMPTOTIC OPTIMALITY OF TRACKING POLICIES IN STOCHASTIC NETWORKS

BY NICOLE BÄUERLE

*University of Ulm*

Control problems in stochastic queuing networks are hard to solve. However, it is well known that the fluid model provides a useful approximation to the stochastic network. We will formulate a general class of control problems in stochastic queuing networks and consider the corresponding fluid optimization problem ( $F$ ) which is a deterministic control problem and often easy to solve. Contrary to previous literature, our cost rate function is rather general. The value function of ( $F$ ) provides an asymptotic lower bound on the value function of the stochastic network under fluid scaling. Moreover, we can construct from the optimal control of ( $F$ ) a so-called tracking policy for the stochastic queuing network which achieves the lower bound as the fluid scaling parameter tends to  $\infty$ . In this case we say that the tracking policy is asymptotically optimal. This statement is true for multiclass queuing networks and admission and routing problems. The convergence is monotone under some convexity assumptions. The tracking policy approach also shows that a given fluid model solution can be attained as a fluid limit of the original discrete model.

**1. Introduction.** There exists a vast literature on optimal control of stochastic queuing networks; see, for example, the recent books by Sennott (1998) and Kitaev and Rykov (1995). Although the dynamic programming technique, which is the most common solution method, is well developed, only a few special networks allow for an explicit solution [cf. also Stidham and Weber (1993)]. Due to the enormous state space of the problems, a numerical solution is often intractable. This has led in recent years to the study of Brownian approximations of the network. See, for example, the survey papers by Harrison (1996) and Williams (1998). These approximations are often more tractable and sometimes a policy can be constructed which is optimal in a heavy traffic regime, that is, when the work load of the system reaches its capacity limit. On the other hand, the class of fluid models has attracted a lot of attention recently, since it has been shown that there is a close connection between the stability of the stochastic network and the corresponding fluid model [see, e.g., Dai (1995), Bramson (1996) and Maglaras (1998)].

Since in examples it often turned out that the optimal control in the fluid problem and the optimal policy in the stochastic network coincide [see, e.g., Bäuerle and Rieder (2000)], the question arises, whether there is also a connection between the control problem in the stochastic network and the fluid problem [cf. Avram, Bertsimas and Ricard (1995), Avram (1997), Atkins and

---

Received May 1999; revised October 1999.

AMS 1991 subject classifications. Primary 60K25; secondary 68M20.

Key words and phrases. Stochastic network, Markov decision process, fluid model, weak convergence, stochastic orderings.

Chen (1995) and Meyn (1997a, 1997b)]. This is an important issue, since several authors have shown that the optimization problem in the fluid setting is often easy to solve. In Avram, Bertsimas and Ricard (1995) one can find numerous scheduling problems which have been solved explicitly using Pontryagin's maximum principle. In addition, the authors give an efficient approximation algorithm to compute the fluid optimal control. In a recent paper, Luo and Bertsimas (1998) developed an algorithm to solve such problems with up to 100 buffers, generalizing previous work by Pullan (1995). Weiss (1995, 1996, 1997) solved several scheduling problems in fluid reentrant lines, showing that modifications of the "last-buffer-first-served" policy are optimal.

In the literature we can find several results concerning the relation between the control problem for stochastic queuing networks and the corresponding fluid problem. Meyn (1997a), for example, proved in Theorem 7.2 that, in the average cost model, the policy iteration, if initialized with a stable policy for the fluid model, yields a sequence of relative value functions which converge when properly normalized against the value function in the fluid model. Chen and Meyn (1998) used this fact to suggest that the value iteration can be sped up when initialized with the value function of the fluid model. In Atkins and Chen (1995) one can find a large numerical study, where the optimal control from the fluid model has been used as a heuristic for the policy in the stochastic network. The performance of this implementation has been compared to simple priority rules. It turned out that the fluid heuristic was often slightly better than the priority rules but not always. Alanyali and Hajek (1998) considered a special routing problem and proved that the load-balancing policy which is optimal in the associate fluid model is asymptotically optimal in the stochastic network. However, the crucial question in general is how to translate the optimal fluid control in an admissible policy for the stochastic network. A direct implementation of the optimal feedback control of  $(F)$  is in general not satisfactory [see, e.g., Kumar and Kumar (1996), Meyn (1997b) and Maglaras (1998)]. For reentrant lines Kumar and Kumar (1996) constructed so-called fluctuation smoothing policies from the fluid optimal control which perform very well. In the general setting, a first proposal came from Meyn (1997b). He used an approach based on an affine shift of the optimal fluid policy which is similar to the requirement of "safety stocks." This idea was subsequently developed in papers by Maglaras (1998, 1999a, b) who used the BIGSTEP idea of Harrison (1996) to construct a class of policies which he called discrete-review policies. These policies are asymptotically optimal under fluid scaling in multiclass queuing networks, that is, when the intensity of the process increases by factor  $\gamma$  and the jump height decreases by the same factor. The idea is to review the state of the system at discrete time points and compute from linear programs the actions that have to be carried out over the next planning period. The information about the fluid model is here put into the LP. Safety stock requirements ensure that the plans can be processed properly.

We will now propose a new class of policies which can be constructed from the optimal fluid control directly and are very intuitive. We will call these

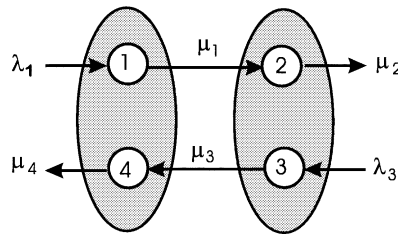


FIG. 1. Rybko–Stolyar network.

policies *tracking policies*. They are asymptotically optimal under fluid scaling in the same sense as in Maglaras (1998) and work for a general class of network problems. The name tracking policy is chosen, since the scaled state process converges to the optimal fluid trajectory. In fact, it is possible to use this approach to track every arbitrary chosen fluid trajectory. Hence, this method works for a large class of objective functions. The approach relies on the observation that in fluid problems the optimal control is usually piecewise constant [see Theorem 3.3 of Pullan (1995) or Theorem 6 of Luo and Bertsimas (1998)]. As a numerical example we have taken the Rybko–Stolyar network in Figure 1 [cf. Rybko and Stolyar (1992) and Maglaras (1998)]: queues 1 and 4 are processed by server 1, while queues 2 and 3 are processed by server 2. The service times of jobs are independent and exponentially distributed with rates  $\mu_1 = \mu_3 = 3$  and  $\mu_2 = \mu_4 = 1.5$ . Queues 1 and 3 receive jobs from outside according to Poisson processes with rate  $\lambda_1 = \lambda_3 = 1$ . The initial state is  $y_0 = (1, 0, 0.5, 1)$  and we assume linear holding cost  $c_1 = \dots = c_n = 1$  for the jobs. The optimization problem is to schedule the servers in order to minimize the expected discounted cost. In Figure 2 we see simulation results of the trajectory of queue 1 under the tracking policy for scaling parameter  $\gamma = 10, 100$  and  $10,000$ . The solid line is the optimal trajectory in the fluid model. In Section 4 we will prove that the trajectories and value functions of the stochastic network under the tracking policy converge against the optimal ones of the fluid model when  $\gamma$  tends to  $\infty$ . Figure 3 shows a simulation result for the trajectory of queues 1–4 respectively, with scaling parameter  $\gamma = 1000$ . The solid line is again the optimal trajectory.

In Table 1 we find the value function  $V_{\sigma^\gamma}^\gamma$  for different scaling parameter  $\gamma$  under the tracking policy  $\sigma^\gamma$ . The optimal cost in the fluid problem is  $7.\bar{2}$ . From the simulation we can see that the tracking policy performs well, when we are close to the limit, that is, in queuing systems with a large initial state and high intensity.

**2. Control problems in stochastic networks.** In this section we present a rather general model for a stochastic queuing network. The state process is formulated as a continuous-time Markov chain  $(Y_t)$  in  $\mathbb{N}_0^N$ , where the  $j$ th component of  $(Y_t)$  gives the number of jobs at queue  $j$  at time  $t$ . The formulation we are using includes admission control, routing, service control

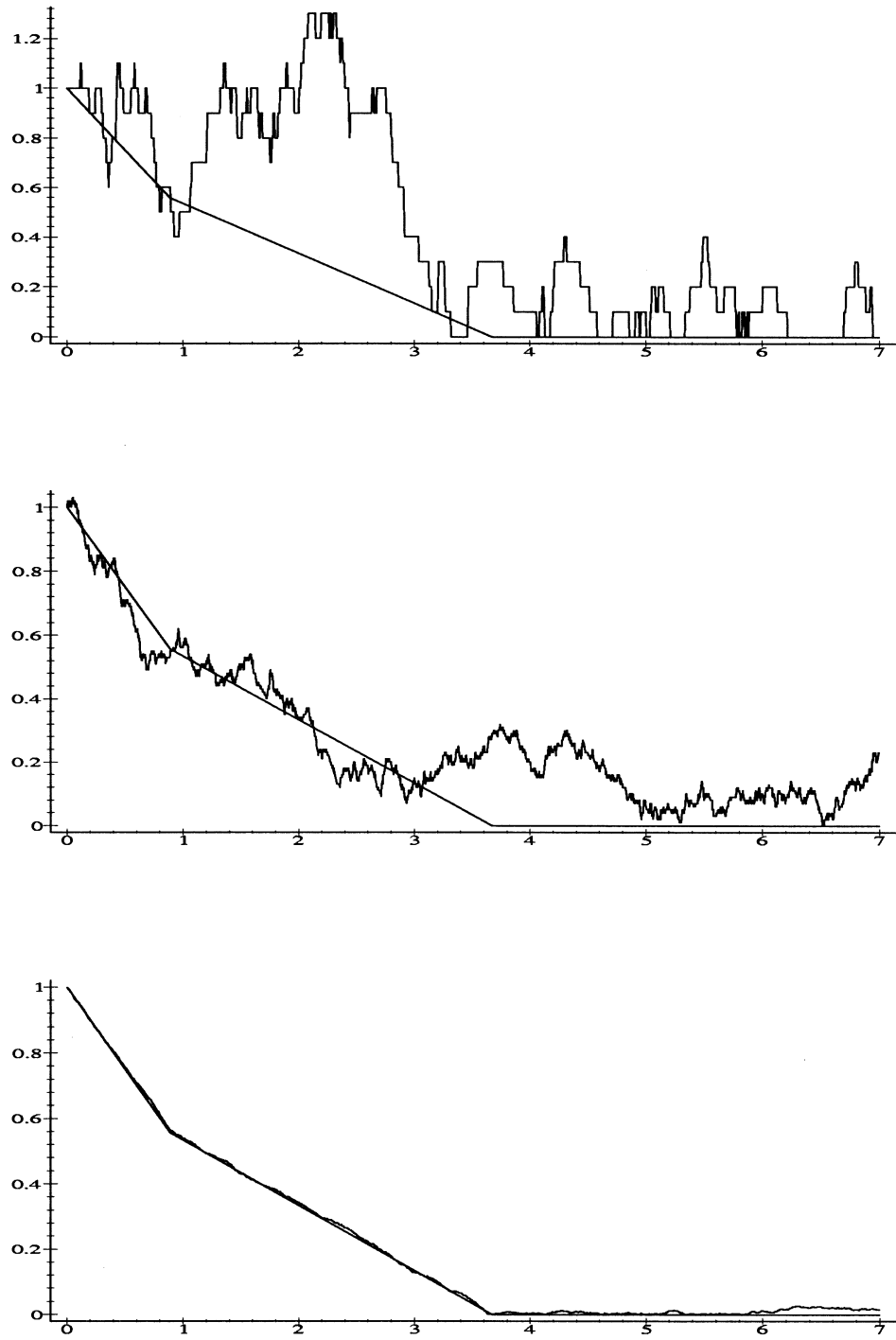


FIG. 2. Trajectories for buffer 1 in the Rybko-Stolyar network with  $\gamma = 10, 10^2, 10^4$ .

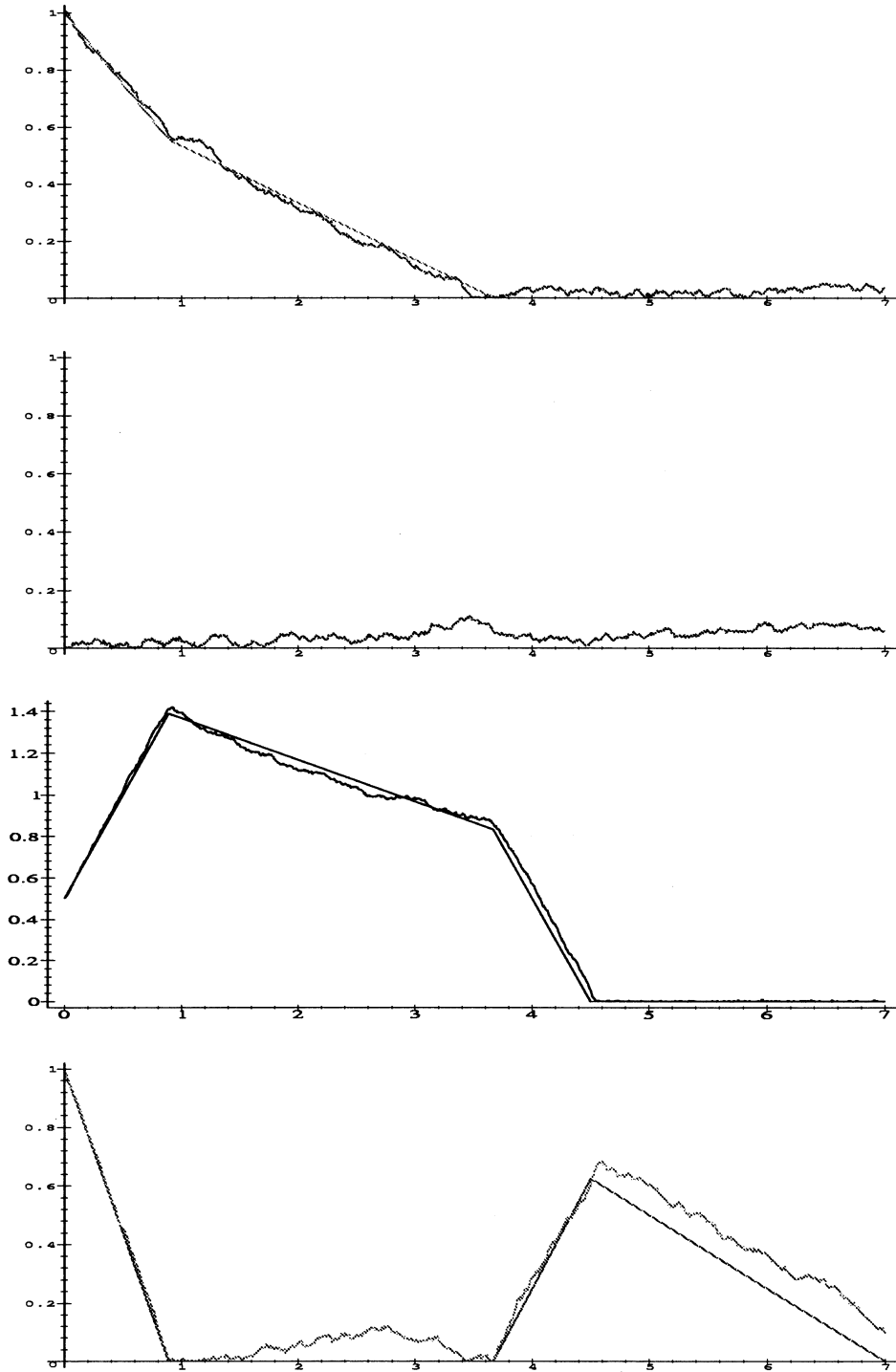


FIG. 3. Trajectories for buffers 1-4 in the Rybko-Stolyar network with  $\gamma = 10^3$ .

TABLE 1  
Value function under tracking policy  $\sigma^\gamma$  for different  $\gamma$

$\gamma$	$10^4$	$10^5$	$10^7$	$\infty$
$V_{\sigma^\gamma}^\gamma(y_0)$	7.5189	7.3522	7.2287	7.2222

and scheduling. To keep the formulation simple, we assume that arrival and service times are independent and exponentially distributed. The model allows us to control the transition rates of the process at each point in time in a nonanticipating fashion. However, it is known that in this case the optimal policy can be found among the discrete-time policies, where decisions have to be taken at state changes only. This leads to the following Markov decision process [see, e.g., Sennott (1998) and Tijms (1986)]: we assume that there are  $N$  queues, and hence the *state space* is  $S = \mathbb{N}_0^N$ . The *action space*  $U \subset \mathbb{R}^K$  has to be compact and convex. The generator  $Q = (q(y, u, y'))$  of  $(Y_t)$  should satisfy the following conditions for all  $y, y' \in S$ :

- (i)  $D(y) := \{u \in U \mid q(y, u, y') = 0 \text{ if } y' \notin S\} \neq \emptyset$ .
- (ii) There exists a linear function  $b : U \rightarrow \mathbb{R}^N$  such that, for all  $u \in D(y)$ ,

$$\sum_{y' \in S} (y' - y)q(y, u, y') = b(u).$$

- (iii) There exists a  $q \in \mathbb{R}_+$  with  $\sup_{u \in U} \sup_{y, y' \in S} |q(y, u, y')| < q$ .

The set  $D(y)$  is the *set of admissible actions* in state  $y$ . As usual, define  $q(y, u) := \sum_{y' \neq y} q(y, u, y')$ .

In the Rybko–Stolyar network of the Introduction we have, for example,  $U = \{u \in [0, 1]^4 \mid u_1 + u_4 \leq 1, u_2 + u_3 \leq 1\}$  and  $D(y) = \{u \in U \mid y_j = 0 \Rightarrow u_j = 0, j = 1, 2, 3, 4\}$ . For  $u \in D(y)$  the generator is

$$\begin{aligned} q(y, u, y + e_1) &= \lambda_1, \\ q(y, u, y + e_3) &= \lambda_3, \\ q(y, u, y - e_2) &= \mu_2 u_2, \\ q(y, u, y - e_4) &= \mu_4 u_4, \\ q(y, u, y + e_2 - e_1) &= \mu_1 u_1, \\ q(y, u, y + e_4 - e_3) &= \mu_3 u_3. \end{aligned}$$

The *cost rate function*  $c : S \times U \rightarrow \mathbb{R}_+$  of the general model should satisfy:

- (i)  $c(y, u) = c_1(y) + c_2(u)$  with  $c_1 : \mathbb{R}^N \rightarrow \mathbb{R}_+$ ,  $c_2 : \mathbb{R}^K \rightarrow \mathbb{R}_+$ ;
- (ii)  $c_1$  is lower semicontinuous,  $c_2$  convex.

Denote by  $(T_n)$ ,  $T_0 := 0$ , the sequence of jump times of the Markov process  $(Y_t)$ . A policy  $\pi = (f_0, f_1, \dots)$  for the Markov decision process is a sequence of

decision rules  $f_n : S \rightarrow U$  with  $f_n(y) \in D(y)$ , where  $f_n$  is applied at time  $T_n$ . For a fixed policy  $\pi$  and initial state  $y \in S$ , there exists a family of probability measures  $P_y^\pi$  on a measurable space  $(\Omega, \mathcal{F})$  and stochastic processes  $(Y_t)$  and  $(\pi_t)$  such that, for  $0 =: T_0 < T_1 < T_2 < \dots$ ,

$$Y_t = Y_{T_n}, \quad T_n \leq t < T_{n+1}, \quad \pi_t = f_n(Y_{T_n}), \quad T_n \leq t < T_{n+1}$$

and

- (i)  $P_y^\pi(Y_0 = y) = P_y^\pi(T_0 = 0) = 1$  for all  $y \in S$ ;
- (ii)  $P_{y_0}^\pi(T_{n+1} - T_n > t | T_0, Y_{T_0}, \dots, T_n, Y_{T_n} = y) = e^{-q(y, f_n(y))t}$  for all  $y \in S, t \geq 0$ ;
- (iii)  $P_{y_0}^\pi(Y_{T_{n+1}} = y' | T_0, Y_{T_0}, \dots, T_n, Y_{T_n} = y, T_{n+1}) = q(y, f_n(y), y')/q(y, f_n(y))$  for  $y, y' \in S, y \neq y'$  and 0, if  $y = y'$ .

We are interested in the discounted cost criterion and define for discount rate  $\beta > 0$

$$V_\pi(y) = E_y^\pi \left[ \int_0^\infty e^{-\beta t} c(Y_t, \pi_t) dt \right].$$

The optimization problem is

$$V(y) = \inf_\pi V_\pi(y).$$

In the remainder of the paper we assume that  $D(y)$  is compact for all  $y \in S$  and the mapping  $u \rightarrow q(y, u, y')$  is continuous for all  $y, y' \in S$ . Under these assumptions, there exists an optimal stationary policy for the  $\beta$ -discounted problem. Moreover, this policy is optimal among all nonanticipating policies. The value iteration is of the form

$$V_{n+1}(y) = \min_{u \in U} \left[ \frac{1}{\beta + q(y, u)} c(y, u) + \frac{1}{\beta + q(y, u)} \left( \sum_{y' \neq y} q(y, u, y') V_n(y') \right) \right].$$

Although problems of this type can, in principle, be solved by policy iteration, the size of the state space, even for simple examples, makes this procedure intractable. Hence, we would be satisfied with a policy which is in some sense “good” and easily computable. Let us now introduce a scaling parameter  $\gamma > 0$  for the stochastic process as follows: let  $\{y^\gamma\}$  be a sequence of initial states such that  $\lim_{\gamma \rightarrow \infty} y^\gamma/\gamma = y$  for  $y \in S$ . To ease notation, we will assume for our problem that  $y^\gamma = \gamma y$  for all  $\gamma \in \mathbb{N}$ , though the proofs are in a more general setting. Denote by  $(\widehat{Y}_t^\gamma)$  the state process with initial state  $y^\gamma$  under a fixed policy  $\pi^\gamma = (f_n^\gamma)$  and define by

$$Y_t^\gamma := \frac{1}{\gamma} \widehat{Y}_{\gamma t}^\gamma$$

the scaled state process. Note that  $(\widehat{Y}_t^\gamma)$  is a process on the state space  $S = \mathbb{N}_0^N$ , whereas  $(Y_t^\gamma)$  is a process on the state space  $(1/\gamma)S$ . If we define the policy  $\tilde{\pi}^\gamma = (\tilde{f}_n^\gamma)$  on the state space  $(1/\gamma)S$  by  $\tilde{f}_n^\gamma((1/\gamma)y) = f_n^\gamma(y)$  and the generator  $\tilde{Q}^\gamma$  by  $\tilde{q}((1/\gamma)y, u, (1/\gamma)y') = \gamma q(y, u, y')$ , then the corresponding process  $(\tilde{Y}_t^\gamma)$  is in distribution equal to the process  $(Y_t^\gamma)$ . The scaled action process is

defined by

$$\pi_t^\gamma := f_n^\gamma(\widehat{Y}_{T_n}^\gamma) \quad \text{if } T_n \leq \gamma t < T_{n+1},$$

where  $(T_n)$  are the jump times of process  $(\widehat{Y}_t^\gamma)$ . As  $\gamma$  tends to  $\infty$ , the intensity of the scaled process increases by factor  $\gamma$ , while the jump heights decrease by the same rate. This scaling is referred to as *fluid scaling*. The scaled value function is then defined by

$$V_{\pi^\gamma}^\gamma(y) = E_y^{\pi^\gamma} \left[ \int_0^\infty e^{-\beta t} c(Y_t^\gamma, \pi_t^\gamma) dt \right].$$

The optimization problem is as before, where we now write  $V_{\pi^\gamma}^\gamma(y)$  and  $V^\gamma(y)$ , respectively, to make the dependence on  $\gamma$  explicit.

Associated with the discounted stochastic network optimization problem is the following deterministic control problem:

$$(F) \begin{cases} \int_0^\infty e^{-\beta t} c(y_t, a_t) dt \rightarrow \min, \\ y_t = y + \int_0^t b(a_s) ds, \\ y_t \geq 0, \\ a_t \in U, \quad t \geq 0. \end{cases}$$

We will call  $(F)$  the *fluid problem*. The value function of this problem will be denoted by  $V^F(y)$  and the optimal control and state trajectory by  $a_t^*$  and  $y_t^*$ , respectively.

**3. An asymptotic lower bound on the value function.** In this section we will show that the value function  $V^F$  of the fluid problem  $(F)$  provides an asymptotic lower bound on the value function  $V_{\pi^\gamma}^\gamma$  of the  $\beta$ -discounted stochastic network, irrespective of the chosen sequence of policies  $(\pi^\gamma)$ . Our approach differs from others, since we have a rather general cost rate function which can also depend on the action. We denote by  $(Y_t^\gamma)$ , for  $\gamma \in \mathbb{N}$ , the state process under fixed policy  $\pi^\gamma = (f_n^\gamma)$  and initial state  $y$ . For the convergence results which follow, the processes  $(Y_t^\gamma)$  are defined on a common probability space  $(\Omega', \mathcal{F}', P_y)$ . Such a probability space can be constructed. As usual, we denote by  $(Y_t^\gamma) \Rightarrow (Y_t)$  the weak convergence of the processes as  $\gamma \rightarrow \infty$ . We understand the processes  $(Y_t^\gamma)$  as random elements with values in  $D^N[0, \infty)$ , which is the space of  $\mathbb{R}^N$ -valued functions on  $[0, \infty)$  that are right continuous and have left-hand limits and are all endowed with the Skorokhod topology. Denote by  $\mathbb{P}(U)$  the set of all probability measures on  $U$  endowed with the Borel  $\sigma$ -algebra. The processes  $(\pi_t^\gamma)$  are random elements of  $\mathcal{A} = \{r : \mathbb{R}_+ \rightarrow \mathbb{P}(U) \mid r \text{ measurable}\}$  which is endowed with the Young topology [cf. Davis (1993)]. It is possible to show that  $\mathcal{A}$  is compact and metrizable. Moreover, measurability and convergence in  $\mathcal{A}$  can be characterized as follows [cf. Rieder (1975)].

LEMMA 1. (a)  $r : \mathbb{R}_+ \rightarrow \mathbb{P}(U)$  is measurable if and only if  $r$  is a transition probability from  $\mathbb{R}_+$  into  $U$ .



(b) Let  $r^n, r \in \mathcal{R}$ .  $r^n \rightarrow r$  for  $n \rightarrow \infty$  if and only if

$$\int_0^\infty \int_U \psi(t, u) r_t^n(du) dt \rightarrow \int_0^\infty \int_U \psi(t, u) r_t(du) dt$$

for all measurable functions  $\psi : \mathbb{R}_+ \times U \rightarrow \mathbb{R}$  such that  $u \mapsto \psi(t, u)$  is continuous for all  $t \geq 0$  and  $\int_0^\infty \sup_{u \in U} |\psi(t, u)| dt < \infty$ .

For the next lemma and Theorem 3 we suppose that  $\pi^\gamma = (f^\gamma)^\infty$  is a stationary policy and define the process

$$M_t^\gamma = Y_t^\gamma - y - \int_0^t b(\pi_s^\gamma) ds.$$

We will first show the following result.

LEMMA 2.  $(M_t^\gamma) \Rightarrow 0$  as  $\gamma \rightarrow \infty$ .

PROOF. Let  $\pi^\gamma = (f^\gamma, f^\gamma, \dots)$  and thus

$$\pi_t^\gamma = f^\gamma(\widehat{Y}_{T_n}^\gamma), \quad T_n \leq \gamma t < T_{n+1}.$$

Denote by  $\mathcal{F}_t^\gamma = \sigma(Y_t^\gamma)$  the  $\sigma$ -algebra generated by the process  $(Y_t^\gamma)$ . From the Dynkin formula we can conclude that  $(M_j^\gamma(t))$ ,  $j = 1, \dots, N$ , is a martingale w.r.t. the filtration  $(\mathcal{F}_t^\gamma)$ . This follows, since, by definition, the generator  $\mathcal{A}$  of the process  $(Y_t^\gamma)$  is

$$\mathcal{A} g\left(\frac{1}{\gamma} y\right) = \sum_{y'} \left( g\left(\frac{1}{\gamma} y'\right) - g\left(\frac{1}{\gamma} y\right) \right) \gamma q(y, f^\gamma(y), y'),$$

where  $g : (1/\gamma)S \rightarrow \mathbb{R}$ . Setting  $g(y) = y_j$ ,  $j = 1, \dots, N$ , it follows with Proposition 14.13 in Davis (1993) and assumption (ii) on the generator that  $(M_j^\gamma(t))$  is a martingale. Define  $\tau_n := \inf\{t \geq 0 \mid M_j^\gamma(t) \geq n\}$ ,  $n \in N$ . Since  $M_j^\gamma$  has jumps of size  $1/\gamma$ ,  $M_j^\gamma(t \wedge \tau_n)$  is bounded and hence a square integrable martingale. Using Fatou's lemma, we obtain

$$\begin{aligned} E_y[(M_j^\gamma(t))^2] &\leq \liminf_{n \rightarrow \infty} E_y[(M_j^\gamma(t \wedge \tau_n))^2] = \liminf_{n \rightarrow \infty} E_y[\langle M_j^\gamma(t \wedge \tau_n) \rangle] \\ &\leq \frac{1}{\gamma^2} E_y[\text{number of jumps in } [0, t]] \leq \frac{1}{\gamma^2} q \gamma t = O\left(\frac{1}{\gamma}\right), \end{aligned}$$

where  $\langle M_j^\gamma(t) \rangle$  is the quadratic variation of  $M_j^\gamma(t)$ . Applying Doob's inequality gives us

$$E_y \left[ \sup_{0 \leq s \leq t} (M_j^\gamma(s))^2 \right] \leq 4 E_y[(M_j^\gamma(t))^2] \leq O\left(\frac{1}{\gamma}\right).$$

Hence we have  $(M_t^\gamma) \Rightarrow 0$  for  $\gamma \rightarrow \infty$  on compact intervals. Applying Theorem VI.16 in Pollard (1984), we obtain  $(M_t^\gamma) \Rightarrow 0$  for  $\gamma \rightarrow \infty$ .  $\square$

**THEOREM 3.** *Every sequence  $(Y_t^\gamma, \pi_t^\gamma)$  has a further subsequence  $(Y_t^{\gamma_n}, \pi_t^{\gamma_n})$  such that  $(Y_t^{\gamma_n}, \pi_t^{\gamma_n}) \Rightarrow (Y_t, R_t)$  and the limit satisfies, with  $\pi_t := \int_U uR_t(du)$*

- (i)  $Y_t = y + \int_0^t b(\pi_s) ds$ ;
- (ii)  $Y_t \in \mathbb{R}_+^N$ ;
- (iii)  $\pi_t \in U$ .

**PROOF.** Let us interpret  $(\pi_t^\gamma)$  as a random element  $(R_t^\gamma) \in \mathcal{A}$ . Hence,  $\pi_t^\gamma = \int_U uR_t^\gamma(du)$  for all  $t \geq 0$ . The first step is to show that the sequence  $(Y_t^\gamma, R_t^\gamma)$  is tight. Due to Proposition 3.2.4 in Ethier and Kurtz (1986), we can do this separately. As far as  $(R_t^\gamma)$  is concerned, it is trivially tight, since  $\mathcal{A}$  is compact. For  $(Y_t^\gamma)$  we use the conditions given in Kushner (1990), Theorem 4.4. That is, we have to check

- (i)  $\lim_{m \rightarrow \infty} \sup_\gamma P_y(\|Y_t^\gamma\| \geq m) = 0$  for all  $t \geq 0$ ;
- (ii)  $\lim_{\delta \rightarrow 0} \limsup_{\gamma \rightarrow \infty} \sup_{\tau \leq T} E_y[\min\{1, \|Y_{\tau+\delta}^\gamma - Y_\tau^\gamma\|\}] = 0$ .

We now make use of the fact that  $(\|Y_t^\gamma - (1/\gamma)y^\gamma\|)$  is stochastically dominated by a Poisson process  $(\Lambda_t^\gamma)$  with parameter  $q\gamma$  and jumps of height  $1/\gamma$ . With the Chebyshev inequality we obtain

$$P_y(\|Y_t^\gamma\| \geq m) \leq \frac{1}{m^2} E_y[\|Y_t^\gamma\|^2] \leq \frac{1}{m^2} \left( (qt)^2 + \frac{qt}{\gamma} + 2qt \frac{\|y^\gamma\|}{\gamma} + \frac{\|y^\gamma\|^2}{\gamma^2} \right),$$

which implies (i). For (ii) we note that  $E_y[\min\{1, \|Y_{\tau+\delta}^\gamma - Y_\tau^\gamma\|\}] \leq \delta q$ . Therefore,  $(Y_t^\gamma, R_t^\gamma)$  is tight, which gives us a subsequence  $(Y_t^{\gamma_n}, R_t^{\gamma_n})$  weakly converging to a limit  $(Y_t, R_t)$ . By Skorokhod's theorem [Ethier and Kurtz (1986), Theorem 3.1.8] the process can be constructed on the same probability space such that the convergence is almost sure. Since  $U$  is convex, we can define  $\pi_t := \int_U uR_t(du) \in U$  for all  $t \geq 0$  and  $\pi_t$  is measurable [cf. Lemma 1(a)]. Using Lemma 1(b), we know that almost surely

$$\int_0^t \int_U uR_s^{\gamma_n}(du) ds \rightarrow \int_0^t \pi_s ds.$$

Together with Lemma 2, (i) and (iii) follow. Because of  $Y_t^\gamma \in \mathbb{R}_+^N$  for all  $\gamma$  we obtain (ii) and the proof is complete.  $\square$

Now we are able to prove the main theorem of this section.

**THEOREM 4.** *For all sequences of policies  $(\pi^\gamma)$  and initial states  $y \in S$ , we obtain*

$$\liminf_{\gamma \rightarrow \infty} V_{\pi^\gamma}^\gamma(y) \geq V^F(y).$$

**PROOF.** Suppose first that  $\pi^\gamma = (f^\gamma, f^\gamma, \dots)$  is a stationary policy. Let  $(Y_t^{\gamma_n}, \pi_t^{\gamma_n})$  be a subsequence such that  $(Y_t^{\gamma_n}, \pi_t^{\gamma_n}) \Rightarrow (Y_t, R_t)$  and  $\pi_t := \int_U uR_t(du)$ ,  $y^{\gamma_n} = \gamma_n y$  for all  $n \in \mathbb{N}$ . Due to the assumption on the

cost function, we have

$$\begin{aligned} E_y \left[ \int_0^\infty e^{-\beta t} c(Y_t^{\gamma_n}, \pi_t^{\gamma_n}) dt \right] &= E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t^{\gamma_n}) dt \right] \\ &\quad + E_y \left[ \int_0^\infty e^{-\beta t} c_2(\pi_t^{\gamma_n}) dt \right] \end{aligned}$$

Let us first look at the second term. Define the mapping  $\hat{c}_2 : \mathcal{R} \rightarrow \mathbb{R}_+$  by

$$\hat{c}_2(r) := \int_0^\infty e^{-\beta t} \int_U c_2(u) r_t(du) dt.$$

It is easy to see that  $\hat{c}_2$  is continuous (cf. Lemma 1) and since  $U$  is compact,  $\hat{c}_2$  is bounded on  $\mathcal{R}$ . Since  $c_2$  is convex we can apply Jensen's inequality and obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} E_y \left[ \int_0^\infty e^{-\beta t} c_2(\pi_t^{\gamma_n}) dt \right] &= E_y \left[ \int_0^\infty e^{-\beta t} \int_U c_2(u) R_t(du) dt \right] \\ &\geq E_y \left[ \int_0^\infty e^{-\beta t} c_2(\pi_t) dt \right]. \end{aligned}$$

Note that we have “=” if  $c_2$  is linear. Now define  $\hat{c}_1^m : D^N[0, \infty) \rightarrow \mathbb{R}_+$  by

$$\hat{c}_1^m(y) := \int_0^m e^{-\beta t} c_1^m(y_t) dt,$$

where  $c_1^m \uparrow c_1$  and  $c_1^m : \mathbb{R}^N \rightarrow \mathbb{R}_+$  is continuous [see Lemma 7.14 in Bertsekas and Shreve (1978)]. Hence,  $\hat{c}_1^m$  is continuous and thus  $\hat{c}_1^m(Y_t^{\gamma_n}) \Rightarrow \hat{c}_1^m(Y_t)$ . Therefore, we obtain with Fatou's lemma and since the convergence  $c_1^m \uparrow c_1$  is monotone

$$\begin{aligned} \liminf_{n \rightarrow \infty} E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t^{\gamma_n}) dt \right] &= \liminf_{n \rightarrow \infty} \lim_{m \rightarrow \infty} E_y [\hat{c}_1^m(Y_t^{\gamma_n})] \\ &\geq \lim_{m \rightarrow \infty} \liminf_{n \rightarrow \infty} E_y [\hat{c}_1^m(Y_t^{\gamma_n})] \\ &\geq \lim_{m \rightarrow \infty} E_y [\hat{c}_1^m(Y_t)] = E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t) dt \right]. \end{aligned}$$

From Theorem 3 we know that the limit  $(Y_t, \pi_t)$  of every converging subsequence is for almost all  $\omega$  an admissible state-action trajectory for the fluid problem  $(F)$ . Hence, we have, in particular,

$$E_y \left[ \int_0^\infty e^{-\beta t} c(Y_t, \pi_t) dt \right] \geq V^F(y)$$

and thus  $\liminf_{\gamma \rightarrow \infty} V_{\pi^\gamma}^\gamma(y) \geq V^F(y)$ . Since for arbitrary policies  $\pi^\gamma = (f_0^\gamma, f_1^\gamma, \dots)$  it holds that  $V_{\pi^\gamma}^\gamma \geq V_{f^\gamma}^\gamma(y)$ , where  $(f^\gamma)^\infty$  is the optimal policy, the statement follows.  $\square$

**4. Asymptotic optimality.** We will show that it is possible at least for some network models to construct a policy in such a way that the lower bound of the last section is achieved in the limit. We will call a policy with this property asymptotically optimal. This notion coincides with the ones used by Meyn (1997a,b) and Maglaras (1998). Note that under Assumption 1 below problem (F) always has an optimal solution. This follows from an existence theorem of Baum [cf. Seierstad and Sydsæter (1987), Theorem 10, page 384]. A crucial observation for this construction is that the optimal control  $a^*$  in problem (F) is often piecewise constant. If, for example, the cost rate function is  $c(y, u) = cy$  which is often the case, this statement follows from Pullan (1995), Theorem 3.3 [cf. Luo and Bertsimas (1998)], since (F) reduces to a *separated continuous linear program*. Otherwise, it is possible to construct for every  $\varepsilon > 0$  a piecewise constant policy which is  $\varepsilon$ -optimal [so-called “chattering theorem”; see, e.g., Kushner and Dupuis (1992), Section 4.6]. The implementation of our policy is a direct translation of the fluid solution. The policy itself is nonstationary; that is, the decision depends also on the current time. A state  $(y, t)$  consists now of the queue length and the time at which the jump occurs. The policy is defined in the following way: suppose that  $a_t^* = u^*(\nu)$  on the interval  $[t_\nu, t_{\nu+1})$ ,  $\nu = 0, 1, \dots, m$ ,  $t_0 := 0$ , and use the decision rule

$$f^\gamma(y, t) = u^*(\nu) \quad \text{if } \gamma t_\nu \leq t < \gamma t_{\nu+1},$$

irrespective of the state  $Y_t$  the network is in. This, of course, may lead to unfeasible allocations where we want to serve a job though there is none there. In such cases we reduce the service rate to 0. We will call a policy of this type a *tracking policy*. Obviously, these policies are nonstationary and the only necessary information about the state is which components are 0. Therefore, this policy is particularly interesting for control problems with no information. Also, it can be implemented in a discrete-review way as explained in Remark 1. We will show that tracking policies are asymptotically optimal for two important classes of control problems in stochastic networks. To do this, we need a further assumption on the cost rate function.

ASSUMPTION 1. (i)  $y \mapsto c_1(y)$  is increasing and convex,  $u \mapsto c_2(u)$  is linear.

(ii) There exist constants  $C_0 \in \mathbb{R}_+$ ,  $k \in \mathbb{N}$  such that, for all  $y \in \mathbb{R}^N$ ,

$$c_1(y) \leq C_0(1 + \|y\|^k).$$

*Multiclass queuing networks* [cf. Dai (1995)]. In the literature the multiclass queuing network is defined as follows: there are  $d$  single-server stations  $k = 1, \dots, d$  and server  $k$  is responsible for the jobs at queue  $j \in K_k \subset \{1, \dots, N\}$ . Each queue  $j$  has exogenous arrivals at rate  $\lambda_j$ . The potential service rate of sever  $k$  is  $\mu_k$ . Upon completion of service of a job at queue  $j$ , it is routed to queue  $i$  with probability  $p_{ji}$ , independent of all previous history. We assume that the routing matrix  $P = (p_{ji})$  is transient, that is,  $\sum_{n=0}^{\infty} P^n < \infty$ . The optimization problem is to schedule the servers among their queues in

order to minimize the discounted expected cost of the system. We obtain this network as a special case of our general model in the following way: denote by  $\{K_1, \dots, K_d\}$ ,  $d < N$ , a partition of the set  $\{1, \dots, N\}$ . The action space is given by  $U = \{u \in [0, 1]^N \mid \sum_{j \in K_k} u_j \leq 1, k = 1, \dots, d\}$ , where  $u_j$  is the fraction of the  $k$ th server that is allocated to queue  $j \in K_k$ . We define the matrix  $A = D(I - P)$ , where  $D$  is an  $N$ -dimensional diagonal matrix with elements  $\mu_j \geq 0$  on the diagonal and  $I$  is the identity matrix. The linear drift function  $b$  is now of the form  $b(u) = \lambda - A^T u$  with  $\lambda \in \mathbb{R}_+^N$ . The set of admissible actions in state  $y \in S$  is  $D(y) = \{u \in U \mid y_j = 0 \Rightarrow u_j = 0, j = 1, \dots, N\}$ . Suppose that  $a_t^*$  is the optimal control in the corresponding fluid model and  $a_t^* = u^*(\nu)$  on  $[t_\nu, t_{\nu+1})$ ,  $\nu = 0, 1, \dots, m$ . The tracking policy  $\sigma^\gamma = (f^\gamma, f^\gamma, \dots)$  is formally defined by

$$f^\gamma(y, t) = u^*(\nu) \wedge \delta(y) \quad \text{if } \gamma t_\nu \leq t < \gamma t_{\nu+1},$$

where  $\wedge$  denotes the componentwise minimum and  $\delta(y) = (\delta_1(y), \dots, \delta_N(y))$  is given by

$$\delta_j(y) = \begin{cases} 0, & \text{if } y_j = 0, \\ 1, & \text{if } y_j > 0. \end{cases}$$

Note that  $f^\gamma(y, t) \in D(y)$  for all  $t \geq 0$ . We will now show the following result.

**THEOREM 5.** *Under Assumption 1, the tracking policy  $\sigma^\gamma$  in the multiclass queuing network satisfies, for  $y \in S$ ,*

$$\lim_{\gamma \rightarrow \infty} V_{\sigma^\gamma}^\gamma(y) = V^F(y),$$

and hence  $\sigma^\gamma$  is asymptotically optimal.

**PROOF.** Let us first consider a continuously defined policy  $\pi_t^\gamma$  with corresponding scaled process  $(Y_t^\gamma)$  which is given by

$$\pi_t^\gamma = u^*(\nu) \wedge \delta(Y_t^\gamma) \quad \text{if } \gamma t_\nu \leq t < \gamma t_{\nu+1}.$$

Denote by  $(\bar{Y}_t^\gamma)$  the scaled process, where we use the tracking policy  $\sigma^\gamma$ . The difference between these two processes is the duration of the time intervals on which the actions  $u^*(\nu) \wedge \delta(y)$  are taken. If  $(T_n^\gamma)$  is the sequence of jump times of process  $(\bar{Y}_t^\gamma)$  and  $N^\gamma(t) := \inf\{n \in \mathbb{N} \mid T_n^\gamma > t\}$ , then we obtain, for  $\gamma \rightarrow \infty$ ,

$$T_{N^\gamma(t)}^\gamma \rightarrow t \quad \text{a.s.}$$

This means that the change points, where we use a different server allocation in the processes  $(Y_t^\gamma)$  and  $(\bar{Y}_t^\gamma)$ , converge together a.s. Hence,  $(Y_t^\gamma)$  and  $(\bar{Y}_t^\gamma)$  have the same limit. Therefore, it suffices to prove the statement for the policy  $\pi_t^\gamma$ . Define  $Y_0^\gamma = y \in S$  for  $\gamma \in \mathbb{N}$ .

On the time interval  $[t_\nu, t_{\nu+1})$  we can think of the process  $(Y_t^\gamma)$  as a Jackson network with  $N$  servers and fixed service rates  $\mu_1 u_1^*(\nu), \dots, \mu_N u_N^*(\nu)$ ,  $\nu = 1, \dots, m$ . In this network server  $k$  is only idle when there is no job at queue  $k$ .

This queuing discipline is called work conserving. We will now look at the process on the time interval  $[0, t_1]$  only. Under the tracking policy we have  $\pi_t^\gamma = U_1^* \hat{\pi}_t^\gamma$ , where  $U_1^* = \text{diag}(u^*(1))$  and  $\hat{\pi}_t^\gamma \in [0, 1]^N$ , and our process fulfills for all  $t \in [0, t_1]$  [ $\mathbb{1}$  denotes the vector  $\mathbb{1} = (1, \dots, 1)$ ]

$$\begin{aligned} (1) \quad & Y_t^\gamma = y + \int_0^t (\lambda - A^T U_1^* \hat{\pi}_s^\gamma) ds - M_t^\gamma \geq 0, \\ (2) \quad & \hat{\pi}_t^\gamma \in [0, 1]^N, \\ (3) \quad & \int_0^\infty Y_t^\gamma (\mathbb{1} - \hat{\pi}_t^\gamma) dt = 0. \end{aligned}$$

As before we can show that every sequence  $(Y_t^\gamma, \hat{\pi}_t^\gamma)$  has a further subsequence  $(Y_t^{\gamma_n}, \hat{\pi}_t^{\gamma_n})$  such that  $(Y_t^{\gamma_n}, \hat{\pi}_t^{\gamma_n}) \Rightarrow (Y_t, \hat{\pi}_t)$  and the limit satisfies, for all  $t \in [0, t_1]$  a.s. [see Dai (1995) for the convergence of (3)],

$$\begin{aligned} (4) \quad & Y_t = y + \int_0^t (\lambda - A^T U_1^* \hat{\pi}_s) ds \geq 0, \\ (5) \quad & \hat{\pi}_t \in [0, 1]^N, \\ (6) \quad & \int_0^\infty Y_t (\mathbb{1} - \hat{\pi}_t) dt = 0. \end{aligned}$$

From Chen [(1995), page 641], we know that the solution  $(Y_t, \hat{\pi}_t)$  of (4)–(6) is unique on the interval  $[0, t_1]$  up to sets of measure 0. However, we know, by definition, that  $u^*(1)$  is admissible for the fluid problem  $(F)$  on  $[0, t_1]$ . Thus, we get that  $(y_t^*, \mathbb{1})$  is the unique solution of (4)–(6) on  $[0, t_1]$ . Since the limit is independent of  $\omega$ , this implies

$$(Y_t^\gamma, \hat{\pi}_t^\gamma) \Rightarrow (y_t^*, \mathbb{1}) \text{ on } [0, t_1].$$

Thus, in particular,  $Y_{t_1}^\gamma \rightarrow y_{t_1}^*$  a.s. Inductively, we obtain in this way that the convergence holds for all  $t \geq 0$ . Now it remains to show that  $\lim_{\gamma \rightarrow \infty} V_{\pi^\gamma}^\gamma(y) = V^F(y)$ . Due to the proof of Theorem 4, it is left to show that

$$\lim_{\gamma \rightarrow \infty} E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t^\gamma) dt \right] = \int_0^\infty e^{-\beta t} c_1(y_t^*) dt.$$

First, since  $c_1 \geq 0$ , it holds that

$$E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t^\gamma) dt \right] = \int_0^\infty e^{-\beta t} E_y [c_1(Y_t^\gamma)] dt.$$

Using Assumption 1(i), we obtain that  $c_1(Y_t^\gamma)$  is stochastically dominated by  $c_1(\Lambda_t^\gamma, \dots, \Lambda_t^\gamma)$  for all  $t \geq 0$ , where  $\Lambda_t^\gamma$  is a Poisson process with parameter  $q\gamma$  and jump heights  $1/\gamma$ . From Bäuerle [(1998), Lemma 1], it follows that, for all  $t \geq 0$  and  $\gamma \geq \gamma'$ ,

$$\Lambda_t^\gamma \leq_{cx} \Lambda_t^{\gamma'},$$

where  $\leq_{cx}$  is the convex ordering. For general  $n$ -dimensional random vectors  $X$  and  $Y$ , we have  $X \leq_{cx} Y$  iff  $E[f(X)] \leq E[f(Y)]$  for all  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  convex [see, e.g., Shaked and Shanthikumar (1994)]. Thus, we obtain, with Assumption 1,

$$E_y[c_1(Y_t^\gamma)] \leq E_y[c_1(\Lambda_t^\gamma, \dots, \Lambda_t^\gamma)] \leq E_y[c_1(\Lambda_t, \dots, \Lambda_t)] < \infty.$$

Moreover, since  $c_1$  is also continuous, we obtain  $c_1(Y_t^\gamma) \Rightarrow c_1(y_t^*)$  for all  $t \geq 0$ . Applying dominated convergence, we obtain

$$\lim_{\gamma \rightarrow \infty} E_y[c_1(Y_t^\gamma)] = c_1(y_t^*).$$

Using Assumption 1(ii), we obtain  $\int_0^\infty e^{-\beta t} E_y[c_1(\Lambda_t, \dots, \Lambda_t)] dt < \infty$ , and applying again dominated convergence yields

$$\lim_{\gamma \rightarrow \infty} E_y \left[ \int_0^\infty e^{-\beta t} c_1(Y_t^\gamma) dt \right] = \int_0^\infty e^{-\beta t} c_1(y_t^*) dt,$$

and the statement is proven.  $\square$

*Admission and routing problems.* Under an admission and routing problem, we understand the following model: there are  $d$  external streams of jobs arriving with intensity  $\lambda_k$ ,  $k = 1, \dots, d$ , and jobs of type  $k$  can be routed to the queues  $j \in K_k \subset \{1, \dots, N\}$ . Each queue  $j$  has a server with potential service rate  $\mu_j$ . The optimization problem is to decide upon admission/rejection of jobs and, in the case of admission, to decide upon the routing of the jobs in order to minimize the discounted expected cost of the system. Our general model specializes to an admission and routing problem in the following way: let  $K_1, \dots, K_d$ ,  $d < N$ , be subsets of the set  $\{1, \dots, N\}$ . The action space is given by  $U = \{(u, v) \in [0, 1]^{d \times N} \times [0, 1]^N \mid u_{kj} = 0 \text{ if } j \notin K_k, \sum_{j \in K_k} u_{kj} \leq 1, k = 1, \dots, d, 0 \leq v_j \leq 1, j = 1, \dots, N\}$ , where  $u_{kj}$  is the fraction of jobs of type  $k$  that is routed to queue  $j$ ;  $v_j$  is the activation level of server  $j$ . Let  $\lambda \in \mathbb{R}_+^d$  and let  $D$  be an  $N$ -dimensional diagonal matrix with elements  $\mu_j \geq 0$  on the diagonal. Thus, the linear function  $b$  is of the form  $b(u) = \lambda u - Dv$ . The set of admissible actions in state  $y \in S$  is  $D(y) = \{(u, v) \in U \mid y_j = 0 \Rightarrow (\lambda u - Dv)_j \geq 0, j = 1, \dots, N\}$ . Suppose that  $\alpha_t^*$  is the optimal control in the corresponding fluid model and  $\alpha_t^* = u^*(v)$  on  $[t_\nu, t_{\nu+1})$ ,  $\nu = 0, 1, \dots, m$ . The tracking policy is exactly defined as before. Hence we obtain the following result.

**THEOREM 6.** *Suppose Assumption 1 is valid.*

(a) *The tracking policy  $\sigma^\gamma$  in the admission and routing problem satisfies, for  $y \in S$ ,*

$$\lim_{\gamma \rightarrow \infty} V_{\sigma^\gamma}^\gamma(y) = V^F(y),$$

*and hence  $\sigma^\gamma$  is asymptotically optimal.*

(b)  *$V_{\sigma^\gamma}^\gamma(y)$  is decreasing in  $\gamma$  for all  $y \in S$ . In particular,  $V^F(y)$  is a lower bound for all  $V_{\sigma^\gamma}^\gamma(y)$ .*

PROOF. (a) The idea is the same as in the proof of Theorem 5. Let us first look at the time interval  $[0, t_1)$ . Since there is no rerouting, each queue separately is an M/M/1-queue with input rates  $\lambda u$  and output rates  $Dv$ . Thus, we obtain

$$Y_t^\gamma \Rightarrow y + (\lambda u^*(1) - Dv^*(1))t$$

on  $[0, t_1)$ . Using the same arguments as before, we can complete the first part of the proof.

(b) Denote by  $\xi_j^\gamma(t) = A_j^\gamma(t) - B_j^\gamma(t)$ ,  $j = 1, \dots, N$ , the difference between a Poisson process  $A_j^\gamma(t)$  with parameter  $\gamma \sum_k \lambda_k u_{kj}$  and jump heights  $1/\gamma$  and a Poisson process  $B_j^\gamma(t)$  with parameter  $\gamma \mu_j$  and jump heights  $1/\gamma$ . The processes  $A_j^\gamma(t)$  and  $B_j^\gamma(t)$  are independent, whereas the processes  $A_1^\gamma(t), \dots, A_n^\gamma(t)$  are not. From Bäuerle (1998) it can be deduced that, for all  $\gamma \geq \gamma'$  and  $0 \leq t_1 < t_2 < \dots < t_n < \infty$ ,

$$\begin{aligned} & (\xi_1^\gamma(t_1), \dots, \xi_1^\gamma(t_n), \dots, \xi_N^\gamma(t_1), \dots, \xi_N^\gamma(t_n)) \\ & \leq_{cx} (\xi_1^{\gamma'}(t_1), \dots, \xi_1^{\gamma'}(t_n), \dots, \xi_N^{\gamma'}(t_1), \dots, \xi_N^{\gamma'}(t_n)), \end{aligned}$$

where  $\leq_{cx}$  denotes the convex ordering. Now it holds that

$$Y_j^\gamma(t) = y_j + \xi_j^\gamma(t) + \sup_{0 \leq s \leq t} (-\xi_j^\gamma(s)).$$

Since this is a convex functional, we obtain, for all  $\gamma \geq \gamma'$  and  $0 \leq t_1 < t_2 < \dots < t_n < \infty$ ,

$$\begin{aligned} & (Y_1^\gamma(t_1), \dots, Y_1^\gamma(t_n), \dots, Y_N^\gamma(t_1), \dots, Y_N^\gamma(t_n)) \\ & \leq_{icx} (Y_1^{\gamma'}(t_1), \dots, Y_1^{\gamma'}(t_n), \dots, Y_N^{\gamma'}(t_1), \dots, Y_N^{\gamma'}(t_n)), \end{aligned}$$

where  $\leq_{icx}$  denotes the increasing convex ordering; that is, for two random vectors it holds that  $X \leq_{icx} Y$  iff  $E[f(X)] \leq E[f(Y)]$  for all  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  increasing, convex [see, e.g., Shaked and Shanthikumar (1994)]. Using the assumptions on  $c_1$ , we obtain

$$\hat{c}_1(Y_t^\gamma) \leq_{icx} \hat{c}_1(Y_t^{\gamma'})$$

for  $\gamma \geq \gamma'$ , and the statement follows.  $\square$

COROLLARY 7. *In the multiclass queuing network, as well as in the admission and routing problem, we have, for  $y \in S$  under Assumption 1,*

$$\lim_{\gamma \rightarrow \infty} V^\gamma(y) = V^F(y).$$

PROOF. From the previous theorems we obtain

$$\begin{aligned} V^F(y) &= \limsup_{\gamma \rightarrow \infty} V_{\sigma^\gamma}^\gamma(y) \geq \limsup_{\gamma \rightarrow \infty} V^\gamma(y) \\ &= \limsup_{\gamma \rightarrow \infty} V_{\hat{\pi}^\gamma}^\gamma(y) \geq \liminf_{\gamma \rightarrow \infty} V_{\hat{\pi}^\gamma}^\gamma(y) \geq V^F(y), \end{aligned}$$



where  $\hat{\pi}^\gamma$  is the optimal policy for scaling parameter  $\gamma$ , which exists due to our assumption, and the proof is complete.  $\square$

REMARK 1. Theorems 5 and 6 can be extended to the case where the interarrival times and service times are i.i.d. but arbitrary [cf. Dai (1995)].

REMARK 2. If the cost rate function satisfies  $c((1/\gamma)y, u) = (1/\gamma)c(y, u)$ , then the value function  $V_\pi^\gamma$  can be expressed with the help of the original value function  $V_\pi$ . An easy substitution gives us

$$V_\pi^\gamma(y) = \frac{1}{\gamma^2} V_\pi^{\beta/\gamma}(\gamma y),$$

where  $V_\pi^{\beta/\gamma}$  is the original value function ( $\gamma = 1$ ) with interest rate  $\beta/\gamma$ .

REMARK 3. tracking policies do not necessarily have to be implemented in an open-loop fashion. It is also possible to stop and review the state of the system after a certain time  $l(|y_0|)$ , where  $l$  is a concave function which tends faster to  $\infty$  than log, but slower than linear [cf. Maglaras (1998)]. Given the new state  $y_1$ , we compute the next tracking policy until time  $l(|y_1|)$  and so on.

REMARK 4. There are several alternatives for the implementation of the tracking policy. We explain the procedures here in the setting of the multiclass queuing network. The only thing one has to make sure is that the fraction of the server allocation to buffer  $j$  is in the long run equal to  $u_j^*(\nu)$  on the time interval  $[t_\nu, t_{\nu+1})$ . If we are not allowed to split the server, there are two possibilities:

(i) We interpret  $u_j^*(\nu)$  as a *randomized decision*; that is, we do a random experiment for each buffer independent of the history, where  $u_j^*(\nu)$  is the probability that the  $k$ th server is assigned to queue  $j \in K_k$ .

(ii) When we can write  $u_j^*(\nu) = \alpha_j / \sum_{i \in K_k} \alpha_i$ , with  $\alpha_j \in \mathbb{N}_0$ ,  $j = 1, \dots, N$ , then we can follow a so-called *generalized round-robin* policy [cf. Dai (1998)]: assign the  $k$ th server in a cyclic fashion  $\alpha_{j_1}$  times to queue  $j_1 \in K_k$ , then  $\alpha_{j_2}$  times to queue  $j_2 \in K_k$  and so on.

**5. Conclusion.** After the stimulating paper by Meyn (1997b) there have been some discussions about the way the optimal fluid control should be translated into the discrete problem. Difficulties arise since the boundary behavior of the fluid model cannot be translated in a one-to-one fashion. The class of tracking policies we have proposed in this paper have the following advantages:

(i) Practically every fluid model solution can be attained as a fluid limit under a tracking policy. Thus, this approach is useful for a large variety of objective functions and also for constrained optimization problems.

(ii) There are different alternatives for the implementation. It is, in particular, useful for control problems with no information.

Of course, tracking policies do not always perform well when implemented; however, they are useful when we are close to the fluid limit. This situation occurs when the initial state is large and the system is operating with high intensity.

**Acknowledgment.** The author thanks an anonymous referee for helpful remarks.

## REFERENCES

- ALANYALI, M. and HAJEK, B. (1998). Analysis of simple algorithms for dynamic load balancing. *Math. Oper. Res.* **22** 840–871.
- ATKINS, D. and CHEN, H. (1995). Performance evaluation of scheduling control of queueing networks: fluid model heuristics. *Queueing Sys.* **21** 391–413.
- AVRAM, F. (1997). Optimal control of fluid limits of queueing networks and stochasticity corrections. In *Mathematics of Stochastic Manufacturing Systems. Lectures in Applied Mathematics* (G. G. Yin and Q. Zhang, eds.) **33** 1–36. Amer. Math. Soc., Providence, RI.
- 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. P. Kelly and R. J. Williams, eds.) 199–234. Springer, New York.
- BÄUERLE, N. (1998). The advantage of small machines in a stochastic fluid production process. *Math. Methods Oper. Res.* **47** 83–97.
- BÄUERLE, N. and RIEDER, U. (2000). Optimal control of single-server fluid networks. *Queueing Systems* **35** 185–200.
- BERTSEKAS, D. P. and SHREVE, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York.
- BRAMSON, M. (1996). Convergence to equilibria for fluid models of FIFO queueing networks. *Queueing Systems* **22** 5–45.
- CHEN, H. (1995). Fluid approximations and stability of multiclass queueing networks: work-conserving disciplines. *Ann. Appl. Probab.* **5** 637–665.
- CHEN, R. R. and MEYN, S. (1998). Value iteration and optimization of multiclass queueing networks. *Queueing Systems* **32** 65–97.
- DAI, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Probab.* **5** 49–77.
- DAI, J. G. (1998). Stability of fluid and stochastic processing networks. Lecture note, Univ. Aarhus.
- DAVIS, M. H. A. (1993). *Markov Models and Optimization*. Chapman and Hall, London.
- ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*. Wiley, New York.
- HARRISON, J. M. (1996). The BIGSTEP approach to flow management in stochastic processing networks. In *Stochastic Networks: Stochastic Control Theory and Applications. Royal Statistical Society Lecture Notes Series* (F. Kelly, S. Zachary and I. Ziedins, eds.) 57–90. Clarendon, Oxford.
- KITAEV, M. and RYKOV, V. (1995). *Controlled Queueing Systems*. CRC Press, Boca Raton, FL.
- KUMAR, S. and KUMAR, P. R. (1996). Fluctuation smoothing policies are stable for stochastic re-entrant lines. *J. Discrete Event Dynam. Sys. Theory Appl.* **6** 361–370.
- KUSHNER, H. J. (1990). Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optim.* **5** 999–1048.
- KUSHNER, H. J. and DUPUIS, P. G. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, New York.
- LUO, X. and BERTSIMAS, D. (1998). A new algorithm for state-constrained separated continuous linear programs. *SIAM J. Control Optim.* **37** 177–210.
- MAGLARAS, C. (1998). Discrete-review policies for scheduling stochastic networks: fluid-scale asymptotic optimality. Preprint.

- MAGLARAS, C. (1999a). Dynamic scheduling in multiclass queueing networks: stability under discrete-review policies. *Queueing Systems* **31** 171–206.
- MAGLARAS, C. (1999b). Discrete-review policies for scheduling stochastic networks: trajectory tracking and fluid-scale asymptotic optimality. Preprint.
- MEYN, S. P. (1997a). The policy iteration algorithm for average reward Markov decision processes with general state space. *IEEE Trans. Automat. Control* **42** 1663–1679.
- MEYN, S. P. (1997b). Stability and optimization of queueing networks and their fluid models. In *Mathematics of Stochastic Manufacturing Systems. Lectures in Applied Mathematics* (G. G. Yin and Q. Zhang, eds.) **33** 175–199. Amer. Math. Soc., Providence, RI.
- POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.
- PULLAN, M. C. (1995). Forms of optimal solutions for separated continuous linear programs. *SIAM J. Control Optim.* **33** 1952–1977.
- RIEDER, U. (1975). Bayesian dynamic programming. *Adv. in Appl. Probab.* **7** 330–348.
- RYBKO, A. N. and STOLYAR, A. L. (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems Inform. Transmission* **28** 199–220.
- SEIERSTAD, A. and SYDSÆTER, K. (1987). *Optimal Control Theory with Economic Applications*. North-Holland, Amsterdam.
- SENNOTT, L. I. (1998). *Stochastic Dynamic Programming and the Control of Queues*. Wiley, New York.
- SHAKED, M. and SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic, New York.
- STIDHAM, S. and WEBER, R. (1993). A survey of Markov decision models for control of networks of queues. *Queueing Systems* **13** 291–314.
- TIJMS, H. C. (1986). *Stochastic Modelling and Analysis: A Computational Approach*. Wiley, Chichester.
- WEISS, G. (1995). On optimal draining of re-entrant fluid lines. In *Stochastic Networks* (F. P. Kelly and R. J. Williams, eds.) 91–103. Springer, New York.
- WEISS, G. (1996). Optimal draining of fluid re-entrant lines: some solved examples. In *Stochastic Networks: Stochastic Control Theory and Applications. Royal Statistical Society Lecture Notes Series* (F. Kelly, S. Zachary, and I. Ziedins, eds.) 19–34. Clarendon, Oxford.
- WEISS, G. (1997). Algorithm for minimum wait draining of two-station fluid re-entrant line. Technical report, Dept. Statistics, Univ. Haifa.
- WILLIAMS, R. J. (1998). Some recent developments for queueing networks. *Probability towards 2000. Lecture Notes in Statist.* 340–356. Springer, New York.

DEPARTMENT OF MATHEMATICS VII  
UNIVERSITY OF ULM  
D-89069 ULM  
GERMANY  
E-MAIL: baeuerle@mathematik.uni-ulm.de