ASYMPTOTIC OPTIMALITY OF TRACKING-POLICIES IN STOCHASTIC NETWORKS

Nicole Bäuerle

University of Ulm

Abstract

Control problems in stochastic queueing 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 queueing networks and consider the corresponding fluid optimization problem (F) which is a deterministic control problem and often easy to solve. On the 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 queueing network which achieves the lower bound as the fluid scaling parameter tends to infinity. In this case we say that the Tracking-policy is asymptotically optimal. This statement is true for multiclass queueing 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.

AMS 1991 subject classification: Primary 60K25, Secondary 68M20

Key words: Stochastic network; Markov decision process; Fluid model; Weak convergence; Stochastic orderings

Short title: Tracking-policies for networks

1 Introduction

There exists a vast literature on optimal control of stochastic queueing networks, see e.g. the recent books by Sennott (1998) and Kitaev/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/Weber (1993)). Due to the enormous state space of the problems, a numerical solution is often intractable. This has led in recent years to study a Brownian approximation of the network. See e.g. 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, i.e. 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), Magalras (1998a)).

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/Rieder (1999)), the question arises, whether there is also a connection between the control problem in the stochastic network and the fluid problem (cf. Avram et al. (1995), Avram (1997), Atkins/Chen (1995), Meyn (1997a,b)). 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 et al. (1995) one can find numerous scheduling problems which have been solved explicitely using Pontryagins maximum principle. In addition, the authors give an efficient approximation algorithm to compute the fluid optimal control. In a recent paper, Luo/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 re-entrant 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 queueing 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/Meyn (1998) used this fact to suggest that the value iteration can be speed up when initialized with the value function of the fluid model. In Atkins/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/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/Kumar (1996), Meyn (1997b), Maglaras (1998)). For re-entrant lines Kumar/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 upon 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 queueing networks, i.e. when the intensity of the process increases by factor γ 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 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/Bertsimas (1998)). As a numerical example we have taken the Rybko-Stolyar network in figure 1 (cf. Rybko/Stolyar (1992), Maglaras (1998)):

queue 1 and 4 are processed by server 1, while queue 2 and 3 are processed by server 2. The service times of jobs are independent and exponentially distributed with rate $\mu_1 = \mu_3 = 3$ and $\mu_2 = \mu_4 = 1.5$. Queue 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 $c_1 = \ldots = c_n = 1$ 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 10000. 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 γ tends to ∞. Figure 3 shows a simulation result for the trajectory of queue 1-4 respectively, with scaling parameter $\gamma = 1000$. The solid line is again the optimal trajectory.
In the next table we find the value function $V^{\gamma}_{\sigma^\gamma}$ for different scaling parameter γ under the Tracking-policy σ^γ. The optimal cost in the fluid problem are $7.\overline{2}$. From the simulation we can see that the Tracking-policy performs well, when we are close to the limit, i.e. in queueing systems with a large initial state and high intensity.

γ	10^4	10^5	10^7	∞
$V^{\gamma}_{\sigma^\gamma}(y_0)$	7.5189	7.3522	7.2287	7.2222

tab.1: Value function under Tracking-policy σ^γ for different γ

2 Control Problems in Stochastic Networks

In this section we present a rather general model for a stochastic queueing 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 and scheduling. To keep the formulation simple we assume that arrival and service times are independent and exponentially distributed. The model allows to control the transition rates of the process at each point in time in a non-anticipating 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), Tijms (1986)): we assume that there are N queues, 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 $b : U \to I\!R^N$ such that for all $u \in D(y)$

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

(iii) there exists a $q \in I\!R_+$ 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* $c : S \times U \to I\!R_+$ of the general model should satisfy

(i) $c(y,u) = c_1(y) + c_2(u)$ with $c_1 : I\!R^N \to I\!R_+,\ c_2 : I\!R^K \to I\!R_+$.

(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 $\pi = (f_0, f_1, \ldots)$ for the Markov decision process is a sequence of decision rules $f_n : S \to U$ with $f_n(y) \in D(y)$, where f_n is applied at time T_n. For a fixed policy π and initial state $y \in S$, there exists a family of probability measures P_y^π on a measurable space $(\Omega, \mathcal{F})$ and stochastic processes (Y_t) and (π_t) such that for $0 =: T_0 < T_1 < T_2 < \ldots$

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

and

(i) $P_y^\pi(Y_0 = y) = P_y^\pi(T_0 = 0) = 1$ for all $y \in S$.

(ii) $P_{y_0}^\pi(T_{n+1} - T_n > t \mid T_0, Y_{T_0}, \ldots, T_n, Y_{T_n} = y) = e^{-q(y, f_n(y))t}$ for all $y \in S, t \geq 0$.

(iii) $P^{\pi}_{y_0}(Y_{T_{n+1}} = y' \mid T_0, Y_{T_0}, \dots, T_n, Y_{T_n} = y, T_{n+1}) = \frac{q(y, f_n(y), y')}{q(y, f_n(y))}$ for $y, y' \in S, y \neq y'$ and zero, if $y = y'$.

We are interested in the discounted cost criterion and define

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

The *optimization problem* is

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

In the sequel we assume that $D(y)$ is compact for all $y \in S$ and the mapping $u \to q(y, u, y')$ is continuous for all $y, y' \in S$. Under these assumptions, there exists an optimal stationary policy for the β-discounted problem. Moreover, this policy is optimal among all non-anticipating 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 $\lim_{\gamma \to \infty} \frac{y^\gamma}{\gamma} = y$ for $y \in S$. To ease notation we will assume for our problem that $y^\gamma = \gamma y$ for all $\gamma \in I\!N$, though the proofs are in a more general setting. Denote by $(\hat{Y}^\gamma_t)$ the state process with initial state y^γ under a fixed policy $\pi^\gamma = (f^\gamma_n)$ and define by

$$Y^\gamma_t := \frac{1}{\gamma} \hat{Y}^\gamma_{\gamma t}$$

the scaled state process. Note that $(\hat{Y}^\gamma_t)$ is a process on the state space $S = I\!N_0^N$, whereas (Y^γ_t) is a process on the state space $\frac{1}{\gamma} S$. If we define the policy $\tilde{\pi}^\gamma = (\tilde{f}^\gamma_n)$ on the state space $\frac{1}{\gamma} S$ by $\tilde{f}^\gamma_n(\frac{1}{\gamma} y) = f^\gamma_n(y)$ and the generator $\tilde{Q}^\gamma$ by $\tilde{q}(\frac{1}{\gamma} y, u, \frac{1}{\gamma} y') = \gamma q(y, u, y')$, then the corresponding process $(\tilde{Y}^\gamma_t)$ is in distribution equal to the process (Y^γ_t). The scaled action process is defined by

$$\pi^\gamma_t := f^\gamma_n(\hat{Y}^\gamma_{T_n}), \quad \text{if } T_n \leq \gamma t < T_{n+1}$$

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

$$V^\gamma_{\pi^\gamma}(y) = E^{\pi^\gamma}_y \left[\int_0^\infty e^{-\beta t} c(Y^\gamma_t, \pi^\gamma_t)\, dt \right].$$

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

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

$$(F) \begin{cases} \int\limits_0^\infty e^{-\beta t} c(y_t, a_t)\; dt \to \min \\ y_t = y + \int\limits_0^t b(a_s)\; ds \\ y_t \geq 0 \\ a_t \in U,\; t \geq 0 \end{cases}$$

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 $V^\gamma_{\pi^\gamma}$ of the β-discounted stochastic network, irrespective of the chosen sequence of policies (π^γ). 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^γ_t), for $\gamma \in I\!N$, the state process under fixed policy $\pi^\gamma = (f^\gamma_n)$ and initial state y. For the convergence results which follow, the processes (Y^γ_t) are defined on a common probability space $(\Omega', \mathcal{F}', P_y)$. Such a probability space can be constructed. As usual, we denote by $(Y^\gamma_t) \Rightarrow (Y_t)$ the weak convergence of the processes as $\gamma \to \infty$. We understand the processes (Y^γ_t) as random elements with values in $D^N[0, \infty)$, which is the space of $I\!R^N$-valued functions on $[0, \infty)$ that are right continuous and have left-hand limits and all endowed with the Skorokhod topology. Denote by $I\!P(U)$ the set of all probability measures on U endowed with the Borel-σ-algebra. The processes (π^γ_t) are random elements of $\mathcal{R} = \{r : I\!R_+ \to I\!P(U) \mid r \text{ measurable } \}$ which is endowed with the Young topology (cf. Davis (1993)). It is possible to show that $\mathcal{R}$ is compact and metrizable. Moreover, measurability and convergence in $\mathcal{R}$ can be characterized as follows (cf. Rieder (1975)):

Lemma 1:

a) $r : I\!R_+ \to I\!P(U)$ is measurable if and only if r is a transition probability from $I\!R_+$ into U.

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

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

for all measurable functions $\psi : I\!R_+ \times U \to I\!R$ 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

Lemma 2: $(M_t^\gamma) \Rightarrow 0$ as $\gamma \to \infty$.

Proof: Let $\pi^\gamma = (f^\gamma, f^\gamma, \ldots)$ and thus

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

Denote by $\mathcal{F}_t^\gamma = \sigma(Y_t^\gamma)$ the σ-algebra generated by the process (Y_t^γ). From the Dynkin formula we can conclude that $(M_j^\gamma(t)),\ j = 1, \ldots, N$ 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^γ) is

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

where $g : \frac{1}{\gamma} S \to I\!R$. Setting $g(y) = y_j,\ j = 1, \ldots, N$ 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^γ has jumps of size $\frac{1}{\gamma}$, $M_j^\gamma(t \wedge \tau_n)$ is bounded and hence a square integrable martingale. Using the Lemma of Fatou we obtain

$$E_y \left[(M_j^\gamma(t))^2 \right] \leq \liminf_{n \to \infty} E_y \left[(M_j^\gamma(t \wedge \tau_n))^2 \right] = \liminf_{n \to \infty} E_y \left[< M_j^\gamma(t \wedge \tau_n) > \right]$$

$$\leq \frac{1}{\gamma^2} E_y [\text{ number of jumps in } [0,t]] \leq \frac{1}{\gamma^2} q \gamma t = O(\frac{1}{\gamma})$$

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

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

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

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 $\pi_t := \int_U u R_t(du)$

(i) $Y_t = y + \int_0^t b(\pi_s)\, ds$.

(ii) $Y_t \in I\!R_+^N$.

(iii) $\pi_t \in U$.

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

(i) $\lim_{m\to\infty} \sup_\gamma P_y(\|Y_t^\gamma\| \geq m) = 0$ for all $t \geq 0$.

(ii) $\lim_{\delta\to 0} \limsup_{\gamma\to\infty} \sup_{\tau\leq T} E_y \left[\min\{1, \|Y_{\tau+\delta}^\gamma - Y_\tau^\gamma\|\}\right] = 0$.

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

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

which implies (i). For (ii) we note that $E_y\left[\min\{1, \|Y_{\tau+\delta}^\gamma - Y_\tau^\gamma\|\}\right] \leq \delta q$. Therefore, (Y_t^γ, R_t^γ) 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/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 $\pi_t := \int_U u R_t(du) \in U$ for all $t \geq 0$ and π_t is measurable (cf. Lemma 1 a)). Using Lemma 1 b) we know that almost sure

$$\int_0^t \int_U u R_s^{\gamma_n}(du)\, ds \to \int_0^t \pi_s\, ds.$$

Together with Lemma 2, (i) and (iii) follow. Because of $Y_t^\gamma \in I\!R_+^N$ for all γ 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 (π^γ) and initial states $y \in S$ we obtain

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

Proof: Suppose first that $\pi^\gamma = (f^\gamma, f^\gamma, \ldots)$ 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 $n \in I\!N$. Due to the assumption on the cost function we have

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

Let us first look at the second term. Define the mapping $\hat{c}_2 : \mathcal{R} \to I\!R_+$ 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

$$\lim_{n\to\infty} E_y\left[\int_0^\infty e^{-\beta t} c_2(\pi_t^{\gamma_n})\, dt\right] \geq E_y\left[\int_0^\infty e^{-\beta t} \int_U c_2(u) R_t(du) dt\right] \geq E_y\left[\int_0^\infty e^{-\beta t} c_2(\pi_t)\, dt\right]$$

Note that we have "=" if c_2 is linear. Now define $\hat{c}_1^m : D^N[0,\infty) \to I\!R_+$ 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 $c_1^m : I\!R^N \to I\!R_+$ is continuous (see Lemma 7.14 in Bertsekas/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 the Lemma of Fatou and since the convergence $c_1^m \uparrow c_1$ is monotone

$$\liminf_{n\to\infty} E_y\left[\int_0^\infty e^{-\beta t} c_1(Y_t^{\gamma_n})\, dt\right] = \liminf_{n\to\infty} \lim_{m\to\infty} E_y\left[\hat{c}_1^m(Y_t^{\gamma_n})\right] \geq$$

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

From Theorem 3 we know that the limit (Y_t, π_t) of every converging subsequence is for almost all ω 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 $\liminf_{\gamma\to\infty} V^\gamma_{\pi^\gamma}(y) \geq V^F(y)$. Since for arbitrary policies $\pi^\gamma = (f_0^\gamma, f_1^\gamma, \ldots)$ it holds that $V^\gamma_{\pi^\gamma} \geq V^\gamma_{f^\gamma}(y)$, 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 (1997) 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/Sydsæter (1987) Theorem 10, p. 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/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 ε-optimal (so-called 'Chattering Theorem' see e.g. Kushner/Dupuis (1992) Section 4.6). The implementation of our policy is a direct translation of the fluid solution. The policy itself is instationary, i.e. 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

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

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 zero. We will call a policy of this type *Tracking-policy*. Obviously these policies are instationary and the only necessary information about the state is which components are zero. 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 $y \in \mathbb{R}^N$

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

Multiclass Queueing Networks (cf. Dai (1995))

In the literature the multiclass queueing network is defined as follows: there are d single-server stations $k = 1, \ldots, d$ and server k is responsible for the jobs at queue $j \in K_k \subset \{1, \ldots, N\}$. Each queue j has exogenous arrivals at rate λ_j. The potential service rate of sever k is μ_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, i.e. $\sum_{n=0}^{\infty} P^n < \infty$. The optimization problem is to schedule the severs 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 $\{K_1, \ldots, K_d\}$, $d < N$ a partition of the set $\{1, \ldots, N\}$. The action space is given by $U = \{u \in [0,1]^N \mid \sum_{j \in K_k} u_j \le 1,\ k = 1, \ldots, d\}$, 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 $\mu_j \ge 0$ 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 $D(y) = \{u \in U \mid y_j = 0 \Rightarrow u_j = 0,\ j = 1, \ldots, N\}$. Suppose that a_t^* is the optimal control in the corresponding fluid model and $a_t^* = u^*(\nu)$ on $[t_\nu, t_{\nu+1})$, $\nu = 0, 1, \ldots, m$. The Tracking-policy $\sigma^\gamma = (f^\gamma, f^\gamma, \ldots)$ is formally defined by

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

where $\wedge$ denotes the componentwise minimum and $\delta(y) = (\delta_1(y), \ldots, \delta_N(y))$ 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 $t \ge 0$. We will now show

Theorem 5:

Under Assumption 1, the Tracking-policy σ^γ in the multiclass queueing network satisfies for $y \in S$

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

and hence σ^γ is asymptotically optimal.

Proof: Let us first consider a continuously defined policy π_t^γ with corresponding scaled process (Y_t^γ) which is given by

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

Denote by $(\bar{Y}_t^\gamma)$ the scaled process, where we use the Tracking-policy σ^γ. 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^γ) is the sequence of jump times of process $(\bar{Y}_t^\gamma)$ and $N^\gamma(t) := \inf\{n \in I\!N \mid T_n^\gamma > t\}$ then we obtain for $\gamma \to \infty$

$$T^\gamma_{N^\gamma(t)} \to t \quad \text{a.s.}$$

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

On time interval $[t_\nu, t_{\nu+1})$ we can think of the process (Y_t^γ) as a Jackson-network with N servers and fixed service rates $\mu_1 u_1^*(\nu), \ldots, \mu_N u_N^*(\nu)$, $\nu = 1, \ldots, m$. In this network server k is only idle when there is no job at queue k. This queueing discipline is called work-conserving. We will now look at the process on time interval $[0, t_1)$ only. Under the Tracking-policy we have $\pi_t^\gamma = U_1^* \hat{\pi}_t^\gamma$ where $U_1^* = diag(u^*(1))$ and $\hat{\pi}_t^\gamma \in [0,1]^N$ and our process fulfills for all $t \in [0, t_1)$ ($1\!\!1$ denotes the vector $1\!\!1 = (1, \ldots, 1)$)

$$Y_t^\gamma = y + \int_0^t \left(\lambda - A^T U_1^* \hat{\pi}_s^\gamma\right) \, ds - M_t^\gamma \geq 0 \tag{1}$$

$$\hat{\pi}_t^\gamma \in [0,1]^N \tag{2}$$

$$\int_0^\infty Y_t^\gamma \, (1\!\!1 - \hat{\pi}_t^\gamma) \, dt = 0 \tag{3}$$

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))

$$Y_t = y + \int_0^t \left(\lambda - A^T U_1^* \hat{\pi}_s\right) \, ds \geq 0 \tag{4}$$

$$\hat{\pi}_t \in [0,1]^N \tag{5}$$

$$\int_0^\infty Y_t \, (1\!\!1 - \hat{\pi}_t) \, dt = 0 \tag{6}$$

From Chen (1995) (p. 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 zero. However, we know by definition that $u^*(1)$ is admissible for the fluid problem (F) on $[0, t_1)$. Thus, we get that $(y_t^*, 1\!\!1)$ is the unique solution of (4)-(6) on $[0, t_1)$. Since the limit is independent of ω, this implies

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

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

$$\lim_{\gamma\to\infty} E_y \left[\int_0^\infty e^{-\beta t} c_1(Y^\gamma_t)\, dt \right] = \int_0^\infty e^{-\beta t} c_1(y^*_t)\, dt.$$

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

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

Using Assumption 1 (i) we obtain that $c_1(Y^\gamma_t)$ is stochastically dominated by $c_1(\Lambda^\gamma_t, \ldots, \Lambda^\gamma_t)$ for all $t \geq 0$, where Λ^γ_t is a Poisson process with parameter $q\gamma$ and jump heights $\frac{1}{\gamma}$. From Bäuerle (1998a) Lemma 1 it follows that for all $t \geq 0$ and $\gamma \geq \gamma'$

$$\Lambda^\gamma_t \leq_{cx} \Lambda^{\gamma'}_t,$$

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 $f : I\!R^n \to I\!R$ convex (see e.g. Shaked/Shanthikumar (1994)). Thus, we obtain with Assumption 1

$$E_y \left[c_1(Y^\gamma_t) \right] \leq E_y \left[c_1(\Lambda^\gamma_t, \ldots, \Lambda^\gamma_t) \right] \leq E_y \left[c_1(\Lambda_t, \ldots, \Lambda_t) \right] < \infty$$

Moreover, since c_1 is also continuous, we obtain $c_1(Y^\gamma_t) \Rightarrow c_1(y^*_t)$ for all $t \geq 0$. Applying dominated covergence we obtain

$$\lim_{\gamma\to\infty} E_y[c_1(Y^\gamma_t)] = c_1(y^*_t).$$

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

$$\lim_{\gamma\to\infty} E_y \left[\int_0^\infty e^{-\beta t} c_1(Y^\gamma_t)\, dt \right] = \int_0^\infty e^{-\beta t} c_1(y^*_t)\, dt$$

and the statement is proven. □

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 λ_k, $k = 1, \ldots, d$ and jobs of type k can be routed to the queues $j \in K_k \subset \{1, \ldots N\}$. Each queue j has a server with potential service rate μ_j. The optimization problem is to decide upon admission/rejection of jobs and in 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 $K_1, \ldots, K_d$, $d < N$ be subsets of the set $\{1, \ldots, N\}$. The action space is given by $U = \{(u, v) \in [0,1]^{d \times N} \times [0,1]^N \mid u_{kj} = 0, \text{ if } j \notin K_k, \ \sum_{j \in K_k} u_{kj} \leq 1, \ k = 1, \ldots, d, \ 0 \leq v_j \leq 1, \ j = 1, \ldots, N\}$, where u_{kj} is the fraction of jobs of type k which is routed to queue j. v_j is the activation level of server j. Let $\lambda \in I\!R_+^d$ 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 $D(y) = \{(u, v) \in U \mid y_j = 0 \Rightarrow (\lambda u - Dv)_j \geq 0, \ j = 1, \ldots, N\}$. Suppose that a_t^* is the optimal control in the corresponding fluid model and $a_t^* = u^*(\nu)$ on $[t_\nu, t_{\nu+1})$, $\nu = 0, 1, \ldots, m$. The Tracking-policy is exactly defined as before. Hence we obtain

Theorem 6:

Suppose Assumption 1 is valid.

a) The Tracking-policy σ^γ in the admission and routing problem satisfies for $y \in S$

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

and hence σ^γ is asymptotically optimal.

b) $V_{\sigma^\gamma}^\gamma(y)$ is decreasing in γ 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 time interval $[0, t_1)$. Since there is no rerouting, each queue separately is an M/M/1-queue with input rates λu 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 $\xi_j^\gamma(t) = A_j^\gamma(t) - B_j^\gamma(t)$, $j = 1, \ldots, N$, the difference between a Poisson process $A_j^\gamma(t)$ with parameter $\gamma \sum_k \lambda_k u_{kj}$ and jump heights $\frac{1}{\gamma}$ and a Poisson process $B_j^\gamma(t)$ with parameter $\gamma \mu_j$ and jump heights $\frac{1}{\gamma}$. The processes $A_j^\gamma(t)$ and $B_j^\gamma(t)$ are independent, whereas the processes $A_1^\gamma(t), \ldots, A_n^\gamma(t)$ are not. From Bäuerle (1998a) it can be deduced that for all $\gamma \geq \gamma'$ and $0 \leq t_1 < t_2 < \ldots < t_n < \infty$

$$(\xi_1^\gamma(t_1), \ldots, \xi_1^\gamma(t_n), \ldots, \xi_N^\gamma(t_1), \ldots, \xi_N^\gamma(t_n)) \leq_{cx} (\xi_1^{\gamma'}(t_1), \ldots, \xi_1^{\gamma'}(t_n), \ldots, \xi_N^{\gamma'}(t_1), \ldots, \xi_N^{\gamma'}(t_n))$$

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

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

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

$$(Y_1^\gamma(t_1),\ldots,Y_1^\gamma(t_n),\ldots,Y_N^\gamma(t_1),\ldots,Y_N^\gamma(t_n)) \leq_{icx} (Y_1^{\gamma'}(t_1),\ldots,Y_1^{\gamma'}(t_n),\ldots,Y_N^{\gamma'}(t_1),\ldots,Y_N^{\gamma'}(t_n))$$

where $\leq_{icx}$ denotes the increasing convex ordering, i.e. for two random vectors it holds that $X \leq_{icx} Y$, iff $E[f(X)] \leq E[f(Y)]$ for all $f : I\!\!R^n \to I\!\!R$ increasing, convex (see e.g. Shaked/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 queueing network as well as in the admission and routing problem we have for $y \in S$ under Assumption 1

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

Proof: From the previous theorems we obtain

$$V^F(y) = \limsup_{\gamma\to\infty} V_{\sigma^\gamma}^\gamma(y) \geq \limsup_{\gamma\to\infty} V^\gamma(y) = \limsup_{\gamma\to\infty} V_{\hat{\pi}^\gamma}^\gamma(y) \geq \liminf_{\gamma\to\infty} V_{\hat{\pi}^\gamma}^\gamma(y) \geq V^F(y),$$

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

Remark 1:

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

b) If the cost rate function satisfies $c(\frac{1}{\gamma}y,u) = \frac{1}{\gamma}c(y,u)$, then the value function V_π^γ can be expressed with the help of the original value function V_π. An easy substitution gives us

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

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

c) 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 infinity than log, but slower that linear (cf. Maglaras (1998)). Given the new state y_1, we compute the next tracking policy until time $l(|y_1|)$ and so on.

d) There are several alternatives for the implementation of the Tracking-policy. We explain the procedures here in the setting of the multiclass queueing 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*, i.e. 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 $u_j^*(\nu) = \frac{\alpha_j}{\sum_{i \in K_k} \alpha_i}$, with $\alpha_j \in I\!N_0, j = 1, \dots, N$, then we can follow a so-called *generalized round-robin* policy (cf. Dai (1998)): assign the k-th server in a cyclic fashion α_{j_1}-times to queue $j_1 \in K_k$, then α_{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 would like to thank an anonymous referee for helpful remarks.

REFERENCES

ALANYALI M AND HAJEK B (1998) Analysis of simple algorithms for dynamic load balancing. *Math. Operations Res.* **22** 840-871.

ATKINS D AND CHEN H (1995) Performance evaluation of scheduling control of queueing networks: fluid model heuristics. *Queueing Systems* **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, eds. Yin GG and Zhang Q, American Mathematical Society, Providence **33** 1-36.

AVRAM F, BERTSIMAS D AND RICARD M (1995) Fluid models of sequencing problems in open queueing networks: an optimal control approach. In *Stochastic Networks* IMA Volumes in Mathematics and its Applications, eds. Kelly FP and Williams RJ, Springer-Verlag, New York 199-234.

BÄUERLE N (1998) The advantage of small machines in a stochastic fluid production process. *Math. Meth. Operations Res.* **47**, 83-97.

BÄUERLE N AND RIEDER U (1999) Optimal control of single-server fluid networks. To appear in *Queueing Systems.*

BERTSEKAS DP AND SHREVE SE (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 RR AND MEYN S (1998) Value iteration and optimization of multiclass queueing networks. *Queueing Systems* **32** 65-97.

DAI JG (1995) On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Probab.* **5** 49-77.

DAI JG (1998) Stability of fluid and stochastic processing networks. Lecture Note at the University of Aarhus.

DAVIS MHA (1993) *Markov models and optimization.* Chapman & Hall, London.

Ethier SN and Kurtz TG (1986) *Markov processes.* John Wiley & Sons, New York.

Harrison JM (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, eds. Kelly F, Zachary S and Ziedins I., Clarendon Press, Oxford 57-90.

Kitaev M and Rykov V (1995) *Controlled queueing systems.* CRC-Press, Boca Raton.

Kumar S and Kumar PR (1996) Fluctuation smoothing policies are stable for stochastic re-entrant lines. *Journal of Discrete Event Dynamic Systems: Theory and Applications* **6** 361-370.

Kushner HJ (1990) Numerical methods for stochastic control problems in continuous time. *SIAM J. Control and Optim.* **5** 999-1048.

Kushner HJ and Dupuis PG (1992) *Numerical methods for stochastic control problems in continuous time.* Springer-Verlag, New York.

Luo X and Bertsimas D (1998) A new algorithm for state-constrained separated continuous linear programs. *SIAM J. Control and 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 SP (1997a) The policy iteration algorithm for average reward Markov decision processes with general state space. *IEEE Transactions on Automatic Control* **42** 1663-1679.

Meyn SP (1997b) Stability and optimization of queueing networks and their fluid models. In *Mathematics of Stochastic Manufacturing Systems* Lectures in Applied Mathematics, eds. Yin GG and Zhang Q, American Mathematical Society, Providence Mathem. **33** 175-199.

Pollard D (1984) *Convergence of stochastic processes.* Springer-Verlag, New York.

Pullan MC (1995) Forms of optimal solutions for separated continuous linear programs. *SIAM J. Control Optim.* **33** 1952-1977.

Rieder U (1975) Bayesian dynamic programming .*Advances of Applied Probability* **7** 330-348.

RYBKO AN AND STOLYAR AL (1992) Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems of Information Transmission* **28** 199-220.

SEIERSTAD A AND SYDSÆTER K (1987) *Optimal Control Theory with economic Applications*, North-Holland, Amsterdam.

SENNOTT LI (1998) *Stochastic dynamic programming and the control of queues.* John Wiley & Sons, New York.

SHAKED M AND SHANTHIKUMAR JG (1994) *Stochastic orders and their applications.* Academic Press, 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 HC (1986) *Stochastic modelling and analysis: a computational approach.* John Wiley & Sons, Chichester.

WEISS G (1995) On optimal draining of re-entrant fluid lines, in *Stochastic Networks* IMA Volumes in Mathematics and its Applications, eds. Kelly FP and Williams RJ, Springer-Verlag, New York 91-103.

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, eds. Kelly F, Zachary S and Ziedins I., Clarendon Press, Oxford 19-34.

WEISS G (1997) Algorithm for minimum wait draining of two-station fluid re-entrant line. Technical Report, Department of Statistics, The University of Haifa.

WILLIAMS RJ (1998) Some recent developments for queueing networks. In *Probability towards 2000*, Lecture Notes in Statistics, eds. Accardi L, Heyde CC, Springer-Verlag, New York 340-356.

Department of Mathematics VII, University of Ulm, D-89069 Ulm, Germany
e-mail: baeuerle@mathematik.uni-ulm.de