CONVEX STOCHASTIC FLUID PROGRAMS WITH AVERAGE COST

Nicole Bäuerle
Department of Mathematics VII
University of Ulm
D-89069 Ulm, Germany
e-mail: baeuerle@mathematik.uni-ulm.de
Phone: +49-731-50-23522, FAX: +49-731-50-23499

Abstract

We consider Stochastic Fluid Programs under the average cost criterion. These models have been introduced in Bäuerle (1999) and are of the following type: suppose (Z_t) is a continuous-time Markov chain with finite state space. As long as $Z_t = z$, the dynamics of the system at time t are given by a linear function $b^z(a(\cdot))$, where a is a control we have to choose. A convex cost rate function c is given, depending on the state and the action. We want to control the system in such a way as to minimize the expected average cost. Such models typically appear in production and telecommunication systems. Using a vanishing discount approach and a discretization technique, we show that the relative value function satisfies a HJB equation and derive a verification theorem. Last but not least we apply our results to manufacturing systems and network problems.

Key words: Stochastic fluid programs; Average cost; Vanishing discount approach; Hamilton-Jacobi-Bellman equation; Verification theorem; Manufacturing system; Multiclass queueing network;

1 Introduction

In this paper we investigate so-called *Stochastic Fluid Programs* (SFP) under the average cost criterion. SFP have been introduced in Bäuerle (1999). They consist of an uncontrollable stochastic process and a controllable deterministic drift. The stochastic process is called *environment process* and influences the dynamics of the system. An informal description is as follows. Suppose $S \subset \mathbb{R}^K$ is the state space of the system and $y_0 \in S$ the initial state. The environment process (Z_t) is assumed to be a continuous-time Markov chain with finite state space Z. We denote by (T_n) the sequence of (partially virtual) jump times of the uniformized process (Z_t). As long as $Z_t = z$, the system evolves according to $y_t = y_0 + \int_0^t b^z(a_s)\, ds$, where $a \in A := \{a : \mathbb{R}_+ \to U \mid a \text{ measurable}\}$ is a control and $b^z : U \to S$ is a given linear function. We assume $U \subset \mathbb{R}^N$. Controls $a \in A$ have to be chosen at the time points (T_n) with the restriction that $y_t \in S$ for all $t \geq 0$. Moreover, a cost rate function $c : S \times Z \times U \to \mathbb{R}_+$ is given and we want to minimize the average cost of the system. Models of this type appear in particular in manufacturing and telecommunications (see e.g. Akella/Kumar (1986), Bilecki/Kumar (1988), Presman et al. (1995), Rajagopal (1995), Rajagopal et al. (1995), Sethi et al. (1997, 1998), Bäuerle/Rieder (1999)). An example is a single-machine, single-product manufacturing system with random breakdowns of the machine. In this case (Z_t) determines the production capacity $\lambda(z)$ of the machine. If we have a constant demand rate μ for the product, then given $Z_t = z$, the dynamics of the system is $y_t = y_0 + \int_0^t \lambda(z) a_s - \mu ds$, where $a_s \in U = [0, 1]$ is the production rate we can choose. $S = \mathbb{R}$ in the model with backlog or $S = \mathbb{R}_+$ in a model without backlog. Thus SFPs are a special class of *piecewise deterministic Markov processes* (see e.g. Davis (1993)) with one exception: in our model we allow for constraints on the actions and the process can move along the boundary of the state space. In the literature there are not many papers about the problem of average cost for piecewise deterministic Markov processes or related models. Most of them deal with β-discounted cost functions or as far as the average cost are concerned with special models. One of the first papers in this area is Bilecki/Kumar (1988). However, their manufacturing system is very special and the analysis relies on the fact that they are able to compute the value function explicitely. In Sethi et al. (1997, 1998) we can find more general - but yet specific production models under the average cost criterion. A similar model has been investigated by Veatch/Caramanis (1999) who characterize the optimal control by switching sets. Hordijk/Van der Duyn Schouten (1983) have dealt with the average cost problem for Markov decision drift processes. In contrast to our model they allow to control the stochastic jumps and not the deterministic drift. Our aim now is to derive conditions for the general SFP which imply the validity of a HJB equation and to derive solution methods. As usual, we use a vanishing discount approach. The discounted cost problem has been solved in Bäuerle (1999) and we cite the results which we need here.

In Section 2 we will first introduce the mathematical model and give a definition of the average cost. We approach the problem by looking at the discrete-time problem. The advantage of this procedure is that we can use the results of Schäl (1993) to solve the average cost problem. The assumptions, which we need to establish our main theorem go back to Sennott (1989a) who used them for problems with a countable state space. Thus we can formulate our main theorems (Theorem 4 and 5) in Section 3. Theorem 4 states the validity of an average cost optimality equation in the time-discrete setting. In Theorem 5 a) we show that the relative value function together with the minimal average cost satisfy a HJB equation. Part b) is a verification theorem and part c) states that under some further conditions an average cost optimal feedback rule can be obtained as a limit of β-discounted optimal decision rules. Since the assumptions of the main theorems are hard to verify directly, we give in Section 4 sufficient conditions for them. Mainly these assumptions imply positive Harris-recurrence of the state processes. Finally, in Section 5 we apply our results to manufacturing systems and scheduling problems in open multiclass queueing networks.

2 Definition of Average Cost for Stochastic Fluid Programs

SFP have been introduced in Bäuerle (1999). Here, we will give the specific formulation which we will use for the investigation of the average cost problem. Suppose (Z_t) is an irreducible, continuous-time Markov chain with finite state space Z and generator Q.

(Z_t) is the *environment process* which influences the dynamics of the system. We suppose that (Z_t) is given as a uniformized process, i.e. let $q > \max_{z \in Z} q_z$ and $P = I + \frac{1}{q}Q$. Then (Z_t) can be constructed from a sequence of jump times (T_n), where the random variables $(T_{n+1} - T_n)$, $n \in \mathbb{N}$ are independent and exponentially distributed with parameter q and from a Markov chain (Λ_n) with transition matrix P as follows. Let $\Lambda_0 := Z_0$, $T_0 := 0$ and $t \geq 0$. Then

$$Z_t = \Lambda_n, \quad \text{if } T_n \leq t < T_{n+1}$$

is in distribution equal to a Markov chain with generator Q. Let $S \subset \mathbb{R}^K$. $E := S \times Z$ is called *state space* of the system. A state $x \in E$ is denoted by $x = (y, z)$. At each jump time point T_n, a function of the *action space* $A = \{a : \mathbb{R}_+ \to U \mid a \text{ measurable}\}$ has to be chosen, depending on the state X_{T_n} of the system, where $U \subset \mathbb{R}^N$. For fixed $z \in Z$, the mapping $b^z : U \to S$ gives the *dynamics* of the system. Let $a \in A$, then

$$\phi_t(x, a) := y + \int_0^t b^z(a_s)\, ds$$

together with z is the state of the system at time t under control a, as long as $Z_t = z$. The set of *admissible actions* in state x is given by

$$D(x) := \{a \in A \mid \phi_t(x, a) \in S, \forall t \geq 0\}.$$

$F := \{f : E \to A \mid f \text{ measurable }\}$ is called the set of *decision rules* and $\pi = (f_n)$, where $f_n \in F$ is called a *policy*. In applications it is more convenient to deal with feedback rules. However, in general it is not clear whether the optimal policy can be described as a feedback rule. A feedback rule $\varphi : E \to U$ is a measurable mapping such that the equation

$$\phi_t(x, \varphi) = y + \int_0^t b^z(\varphi(\phi_s(x, \varphi), z)) ds$$

has a unique solution $\phi_t(x, \varphi)$, $t \geq 0$ and $\phi_t(x, \varphi) \in S$ for all $t \geq 0$. Last but not least, a measurable *cost rate function* $c : E \times U \to \mathbb{R}_+$ is given. For a fixed policy $\pi = (f_n)$ there exist a family of probability measures $\{P_x^\pi \mid x \in E\}$ on a measurable space $(\Omega, \mathcal{F})$ and stochastic processes $(X_t) = (Y_t, Z_t)$ and (π_t) such that for $0 = T_0 < T_1 < T_2 < \ldots$

$$(Y_t, Z_t) = (\phi_{t-T_n}(X_{T_n}, f_n(X_{T_n})), Z_{T_n}), \quad \pi_t = f_n(X_{T_n})(t) \ \text{ for } \ T_n \leq t < T_{n+1}$$

and

(i) $P_x^\pi(X_0 = x) = P_x^\pi(T_0 = 0) = 1$ for all $x \in E$.

(ii) $P_x^\pi(T_{n+1} - T_n > t \mid T_0, X_{T_0}, \dots, T_n, X_{T_n}) = e^{-qt}$.

(iii) $P_x^\pi(X_{T_{n+1}} \in B \times \{z'\} \mid T_0, X_{T_0}, \dots, X_{T_n}, T_{n+1}) =$
$p_{Z_{T_n} z'} \, 1_B \left(\phi_{T_{n+1}-T_n}(X_{T_n}, f_n(X_{T_n})(T_{n+1} - T_n))\right)$ for $z' \in Z$ and $B \in \mathfrak{B}(S)$.

If the policy π is stationary, i.e. $\pi = f^\infty$, we write P_x^f, if π can be represented by a feedback rule φ we write P_x^φ. The average cost are now defined in the following way.

Definition 1:

Let a SFP be given and let π be a policy. For $x \in E$ define

a) the *average cost under policy* π*, starting the system in* x by

$$G_\pi(x) = \limsup_{t\to\infty} \frac{1}{t} E_x^\pi \left[\int_0^t c(X_s, \pi_s) \, ds \right].$$

b) the *minimal average cost, starting the system in* x by

$$G(x) = \inf_\pi G_\pi(x).$$

c) π is called *average optimal*, if it attains the infimum in b) for all $x \in E$.

Now suppose a feedback rule φ is given. An important role in the following analysis plays the extended generator $\mathcal{A}_\varphi$ of the state process (X_t). According to Davis (1993) Theorem 26.14 the domain $\mathcal{D}(\mathcal{A}_\varphi)$ is given by (cf. also Rolski et al. (1999) Theorem 11.2.2)

$$\begin{aligned} \mathcal{D}(\mathcal{A}_\varphi) \;\; = \;\; & \{v : E \to I\!\!R \mid v \text{ measurable }, t \mapsto v(\phi_t(x,\varphi), z) \text{ is absolutely} \\ & \text{continuous for all } \; x \in E\} \end{aligned}$$

and a version of the extended generator itself is defined for $v \in \mathcal{D}(\mathcal{A}_\varphi)$ as

$$\mathcal{A}_\varphi v(x) = \lim_{t\to 0} \frac{1}{t}(E_x^\varphi[v(X_t)] - v(x)) = \tilde{v}(x) + q \sum_{z'} p_{zz'}(v(y,z') - v(x))$$

where $\tilde{v} : E \to I\!\!R$ is such that

$$v(\phi_t(x,\varphi), z) - v(x) = \int_0^t \tilde{v}(\phi_s(x,\varphi), z) \, ds.$$

In particular it holds for $v \in \mathcal{D}(\mathcal{A}_\varphi)$ that

$$E_x^\varphi[v(X_t)] - v(x) = E_x^\varphi\left[\int_0^t \mathcal{A}_\varphi v(X_s)ds\right]. \tag{1}$$

For $u \in U(x) := \{u \in U \mid \exists \delta > 0 : \; y + tb^z(u) \in S, \; 0 \le t \le \delta\}$ the generator $\mathcal{A}_u$ is defined by $\mathcal{A}_u v(x) := \mathcal{A}_{\tilde{\varphi}} v(x)$ with $\tilde{\varphi}(y + tb^z(u)) = u$ for $0 \le t \le \delta$. Note that if v is convex, $\mathcal{A}_u v$ is always well-defined, since directional derivatives exist in this case.

In view of the time-discretization which we need in section 3, we define now by

$$C(x,a) := \int_0^\infty e^{-qt} c(\phi_t(x,a), z, a_t)dt$$

the expected cost between two jumps of the environment process, when the state after the last jump is $x \in E$ and action $a \in D(x)$ is taken. The probability

$$p(x,a; B \times \{z'\}) := q\, p_{zz'} \int_0^\infty e^{-qt} 1_B\left(\phi_t(x,a)\right)\; dt$$

gives the one-step probability of getting from state $x \in E$ under action $a \in D(x)$ in a state in $B \times \{z'\}$ after one transition of the environment process.

3 A HJB Equation

In this section we will prove the validity of an average cost optimality equation in discrete time, the validity of a HJB equation and a verification theorem. We use the vanishing discount approach to derive our results. Therefore, we define for an interest rate $\beta > 0, x \in E$

$$V^\beta(x) = \inf_\pi E_x^\pi\left[\int_0^\infty e^{-\beta t} c(X_t, \pi_t)\, dt\right]$$

as the minimal expected discounted cost, starting in x. It has been shown in Bäuerle (1999) that under the following Assumption 1 the *β-discounted cost optimality equation* holds, i.e. for all $x \in E$

$$V^\beta(x) = \min_{a \in D(x)}\left[\int_0^\infty e^{-(\beta+q)t}\left\{c(\phi_t(x,a), z, a_t) + q\sum_{z'} p_{zz'} V^\beta\left(\phi_t(x,a), z'\right)\right\}\, dt\right] \tag{2}$$

and a β-discounted stationary optimal policy exists. (Indeed, the assumptions presented in Bäuerle (1999) are even weaker).

Assumption 1:

(i) $S = I\!R^K$ or $S = I\!R^K_+$ and U is convex and compact w.r.t. the usual Euclidian norm.

(ii) $u \mapsto b^z(u)$ is linear, $(y,u) \mapsto c(y,z,u)$ is convex and continuous for all $z \in Z$.

(iii) For all $\beta > 0$, there exists a policy π^β such that $V^\beta_{\pi^\beta}(x) < \infty$ for all $x \in E$ and if $S = I\!R^K_+$, $V^\beta(y,z)$ is increasing in y.

For fixed $\xi \in E$ we will now define

$$h^\beta(x) = V^\beta(x) - V^\beta(\xi) \quad \text{and} \quad \rho(\beta) = \beta V^\beta(\xi),$$

where h^β is called *relative value function*. Under the following assumptions we will derive a HJB equation and a verification theorem. Assumption 2 has essentially been established by Sennott (1989a) for Markov decision processes with a countable state space.

Assumption 2:

(i) There exists a policy π such that $G_\pi(x) < \infty$ for all $x \in E$.

(ii) There exist constants $L \in I\!R$, $\bar\beta > 0$ and an upper semicontinuous function $M : E \to I\!R_+$ with

$$L \le h^\beta(x) \le M(x)$$

for all $x \in E$ and $0 < \beta \le \bar\beta$ and $\int_E M(x')p(x,a;dx') < \infty$ for all $x \in E,\ a \in D(x)$.

In the sequel we summarize auxiliary statements before proving our main theorems. The following Tauber Theorem will be a useful tool. A version of it can be found e.g. in Hordijk/Van der Duyn Schouten (1983) (Lemma 4.5).

Theorem 1:

For all policies π and $x \in E$ we obtain: $\limsup_{\beta \downarrow 0} \beta V^\beta_\pi(x) \le G_\pi(x)$.

Applying the Tauber Theorem we immediately obtain the following lemma (cf. Hernández-Lerma/Lasserre (1996) sec.5).

Lemma 2:

Suppose Assumption 2 is valid.

a) There exists a sequence of interest rates $\beta_n \downarrow 0$ such that for all $x \in E$

$$0 \leq \lim_{n\to\infty} \beta_n V^{\beta_n}(x) = \limsup_{\beta\downarrow 0} \rho(\beta) < \infty.$$

b) For all policies π and $x \in E$ it holds: $\limsup_{\beta\downarrow 0} \rho(\beta) \leq G_\pi(x)$.

Lemma 3:

Suppose that Assumptions 1 and 2 hold.

a) The relative value functions $h^\beta(y,z)$ are convex in y for all $z \in Z$.

b) Every sequence $\beta_n \downarrow 0$ has a further subsequence $\beta_{n_m} \downarrow 0$ such that $h^{\beta_{n_m}}(x) \to h(x)$ uniform on compact sets and $h(y,z)$ is convex in y for all $z \in Z$.

Proof: Part a) follows from Bäuerle (1999) Lemma 5. For part b) we show that for every sequence $\beta_n \downarrow 0$, the sequence (h^{β_n}) is uniformly locally Lipschitz-continuous, i.e. for $z \in Z$ and every $y, y' \in S$ with $\|y\|, \|y'\| \leq r$, there exists a constant $C = C(z,r)$ independent of β such that

$$|h^\beta(y,z) - h^\beta(y',z)| \leq C\|y - y'\|$$

for all $0 < \beta \leq \bar{\beta}$. This implies that the sequence (h^{β_n}) is equicontinuous. Since $h^\beta(x) \leq M(x)$ for β small enough, the assertion follows with the Arzela-Ascoli Theorem. The convexity of h follows directly from the convexity of h^β, $\beta > 0$.

Now fix $z \in Z$. Suppose first that $S = I\!R^K$ and let $y, y' \in S$ with $\|y\|, \|y'\| \leq r$ and $\varepsilon > 0$. Define

$$\hat{y} := y + \frac{\varepsilon}{\|y - y'\|}(y - y').$$

Then $\|\hat{y}\| \leq \|y\| + \varepsilon \leq r + \varepsilon$. Since M is upper semicontinuous we have that for all $\|\tilde{y}\| \leq r + \varepsilon$ and $0 < \beta \leq \bar{\beta}$:

$$|h^\beta(\tilde{y},z)| \leq \max_{\|y\|\leq r+\varepsilon} M(y,z) =: \hat{C} = \hat{C}(z, r+\varepsilon).$$

Moreover, since

$$y = \frac{\|y-y'\|}{\|y-y'\|+\varepsilon}\,\hat{y} + \frac{\varepsilon}{\|y-y'\|+\varepsilon}\,y'$$

and due to the convexity of h^β we obtain

$$h^\beta(y,z) \le \frac{\|y-y'\|}{\|y-y'\|+\varepsilon}\,h^\beta(\hat{y},z) + \frac{\varepsilon}{\|y-y'\|+\varepsilon}\,h^\beta(y',z)$$

and therefore

$$|h^\beta(y,z)-h^\beta(y',z)| \le \frac{\|y-y'\|}{\|y-y'\|+\varepsilon}\,|h^\beta(\hat{y},z)-h^\beta(y',z)| \le \frac{\|y-y'\|}{\|y-y'\|+\varepsilon}2\hat{C} \le \frac{2\hat{C}}{\varepsilon}\|y-y'\|$$

for all $0<\beta\le\bar{\beta}$.

If $S=\mathbb{R}_+^K$ we proceed as follows. Define an extension of h^β on $\mathbb{R}^K$ as: $\hat{h}^\beta(y,z) := h^\beta(y\vee 0,z)$, where $y\vee y'$ denotes the componentwise maximum of y and $y'\in\mathbb{R}^K$. It obviously holds that $\hat{h}^\beta(y,z)=h^\beta(y,z)$ for all $y\in S$ and $\hat{h}^\beta(y,z)$ is increasing in y due to Assumption 1. The functions $y\mapsto\hat{h}^\beta(y,z)$ are again convex on $\mathbb{R}^K$ since for $y,y'\in\mathbb{R}^K$ and $\lambda\in[0,1]$:

$$\begin{aligned}
\hat{h}^\beta(\lambda y+(1-\lambda)y',z) &\le \hat{h}^\beta(\lambda(y\vee 0)+(1-\lambda)(y'\vee 0),z) = \\
&= h^\beta(\lambda(y\vee 0)+(1-\lambda)(y'\vee 0),z) \le \\
&\le \lambda h^\beta(y\vee 0,z)+(1-\lambda)h^\beta(y'\vee 0,z) = \lambda\hat{h}^\beta(y,z)+(1-\lambda)\hat{h}^\beta(y',z)
\end{aligned}$$

As before we can conclude that $(\hat{h}^\beta)$ is equicontinuous on $\mathbb{R}^K$ which implies that (h^β) is equicontinuous on S and the assertion follows. □

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

Theorem 4: (*Average cost optimality equation*)
Suppose that Assumptions 1 and 2 hold. Then

a) There exists a constant $\rho\ge 0$ and a convex function $h: E\to\mathbb{R}$ such that the *average cost optimality equation* holds, i.e. for all $x\in E$

$$\frac{\rho}{q}+h(x) = \min_{a\in D(x)}\left[C(x,a)+\int_E h(x')p(x,a;dx')\right]. \tag{3}$$

b) There exists a minimizer f^0 of (3) (i.e. $f^0(x)$ attains the minimum on the right-hand-side of (3) for $x \in E$) and a sequence $\beta_m \downarrow 0$ such that

$$f^0(x) = \lim_{m\to\infty} f^{\beta_m}(x),$$

where f^{β_m} is an optimal decision rule in the β_m-discounted model.

Proof: Define $\rho = \limsup_{\beta\downarrow 0} \rho(\beta) \geq 0$ which is finite because of Lemma 2 a). Take β_n as the subsequence such that $\rho = \lim_{n\to\infty} \rho(\beta_n)$. From Lemma 3 we know that there exists a further subsequence (for convenience still denoted by β_n) such that

$$h(x) := \lim_{n\to\infty} h^{\beta_n}(x)$$

uniform on compact sets and h is convex. Using the validity of the β-discounted cost optimality equation we proceed as in Schäl (1993) Theorem 3.8 to obtain

$$\frac{\rho}{q} + h(x) \geq C(x, f^0(x)) + \int_E h(x')p(x, f^0(x); dx') \geq \min_{a\in D(x)} \left[C(x,a) + \int_E h(x')p(x,a; d\,x') \right]$$

where $f^0(x)$ is an accumulation point of a certain sequence $\{f^{\beta_m}(x)\}$ with $\beta_m \downarrow 0$ for $m \to \infty$. On the other hand we have from the discounted optimality equation for all $x \in E, a \in D(x)$

$$\begin{aligned}
\frac{\rho(\beta_n)}{q} + h^{\beta_n}(x) &= \min_{a\in D(x)} \left[\int_0^\infty e^{-(\beta_n+q)t} \left\{ c(\phi_t(x,a), z, a_t) + q \sum_{z'} p_{zz'} h^{\beta_n}\left(\phi_t(x,a), z'\right) \right\} dt \right] \\
&\leq C(x,a) + \int_0^\infty e^{-(\beta_n+q)t} q \sum_{z'} p_{zz'} h^{\beta_n}\left(\phi_t(x,a), z'\right)\, dt
\end{aligned}$$

Taking $n \to \infty$ we obtain with Assumption 2 (ii) and Dominated Convergence

$$\frac{\rho}{q} + h(x) \leq C(x,a) + \int_E h(x')p(x,a; d\,x')$$

for all $x \in E$, $a \in D(x)$. Altogether we have now shown equation (3) and that the decision rule f^0 attains the minimum on the right-hand-side. $\square$

Theorem 5: (*HJB equation*)

Suppose that Assumptions 1 and 2 hold. Then

a) $\rho \geq 0$ and $h : E \to I\!R$ of Theorem 4 satisfy the following HJB equation for all $x \in E$

$$\rho = \min_{u \in U(x)} \left[c(x,u) + \mathcal{A}_u h(x) \right]. \tag{4}$$

b) Every feedback rule φ which satisfies

$$\rho \geq c(x, \varphi(x)) + \mathcal{A}_\varphi h(x)$$

for all $x \in E$ where $t \mapsto h(\phi_t(x,\varphi), z)$ is differentiable at $t = 0$, is average optimal and ρ are the minimal average cost.

c) Suppose f^0 of Theorem 4 b) is given by a feedback rule φ^0 and either $c(x,u)$ is independent of u or the set of discontinuity points of $t \mapsto \varphi^0(y_t^0)$ is of measure zero. Then φ^0 is average optimal and ρ are the minimal average cost.

Proof: For $x \in E, u \in U$ let us define $G(x,u) := c(x,u) + q \sum_{z'} p_{zz'} h(y, z')$. Written in a slightly different form, equation (3) reads

$$\begin{aligned} h(x) &= \min_{a \in D(x)} \left[\int_0^\infty e^{-qt} (G(\phi_t(x,a), a_t) - \rho) dt \right] \\ &= \int_0^\infty e^{-qt} (G(\phi_t(x,\varphi^0), \varphi^0(\phi_t(x,\varphi^0))) - \rho) dt. \end{aligned}$$

With the usual arguments we can show that the Bellman principle holds, i.e. for $T > 0$

$$h(x) = \min_{a \in D(x)} \left[\int_0^T e^{-qt} (G(\phi_t(x,a), a_t) - \rho) dt + e^{-qT} h(\phi_T(x,a), z) \right].$$

Thus, we obtain for $a \in D(x)$ with $a_t \equiv u,\ 0 \leq t \leq T$

$$\frac{1}{T} \left(h(x) - e^{-qT} h(\phi_T(x,a), z) \right) \leq \frac{1}{T} \int_0^T e^{-qt} (G(\phi_t(x,a), u) - \rho) dt.$$

Note that $G(x,u)$ is continuous in y (cf. Lemma 3) and since $h(y,z)$ is convex in y, $\mathcal{A}_u h(x)$ is well-defined. Thus, we obtain with $T \to 0$

$$\rho \leq c(x,u) + \mathcal{A}_u h(x).$$

Therefore, $\rho \leq \min_{u \in U(x)} \left[c(x,u) + \mathcal{A}_u h(x) \right]$. Now suppose for $\varepsilon > 0$, $\rho < c(x,u) + \mathcal{A}_u h(x) + \varepsilon$ for all $u \in U(x)$. Thus for any $a \in D(x)$ due to the continuity of the right-hand-side expression in x (note that the convexity of h implies that the directional derivatives are continuous, see Theorem 25.4 in Rockafellar (1972)), it holds that

$$\rho < c(\phi_t(x,a),a_t) + \mathcal{A}_{a_t} h(\phi_t(x,a),z) + \varepsilon'$$

for t small enough ($t \le T$) and $\varepsilon' > 0$. Thus, we get

$$\int_0^T e^{-qt}(G(\phi_t(x,a),a_t) - \rho)dt + e^{-qT} h(\phi_T(x,a),z) >$$

$$> \int_0^T e^{-qt}(qh(\phi_t(x,a),z) - h'_{a_t}(\phi_t(x,a),z)dt + e^{-qT} h(\phi_T(x,a),z) + \varepsilon'' = h(x) + \varepsilon'',$$

where $h'_u(x) := \lim_{t\downarrow 0} \frac{1}{t}(h(y+tb(u),z) - h(y,z))$ is the directional derivative along direction $b(u)$ and $\varepsilon'' > 0$. The last equation follows from the Dynkin formula for convex functions (cf. Sethi/Zhang (1994), p.74). Taking the infimum over all $a \in D(x)$ gives $h(x) > h(x)$. Hence, our assumption was false and we obtain now $\rho = \min_{u\in U(x)} [c(x,u) + \mathcal{A}_u h(x)]$. For part b) suppose a feedback rule φ is given which satisfies $\rho \ge c(x,\varphi(x)) + \mathcal{A}_\varphi h(x)$. Integrating over t from 0 to T gives us

$$\rho T \ge \int_0^T c(X_t,\varphi(X_t))dt + \int_0^T \mathcal{A}_\varphi h(X_t)dt,$$

where (X_t) is the state process induced by φ. Hence with Assumption 2 and formula (1) we obtain

$$\begin{aligned} \rho &\ge \frac{1}{T} E_x^\varphi \left[\int_0^T c(X_t,\varphi(X_t))dt \right] + \frac{1}{T} E_x^\varphi [h(X_T)] - \frac{1}{T} h(x) \ge \\ &\ge \frac{1}{T} E_x^\varphi \left[\int_0^T c(X_t,\varphi(X_t))dt \right] + \frac{L}{T} - \frac{1}{T} h(x) \end{aligned}$$

Taking $\limsup_{T\to\infty}$ yields $\rho \ge G_\varphi(x)$. Since we always have $\rho \le G_\varphi(x)$ due to Lemma 2 the assertion follows. For part c) we show first in the same way as part a) that $\rho = c(x,u) + \mathcal{A}_{\varphi^0} h(x)$ for almost all $x \in E$ which lie on the path generated by φ^0. Part b) implies then the result. □

4 Verifying the Assumptions

Assumption 2 is often difficult to verify directly. However, we can give some sufficient conditions which will prove extremely useful in our applications. For the next lemma suppose that $c \ge 1$, otherwise replace c by $c+1$.

Lemma 6:

Suppose that Assumption 1 is valid and that there exists a decision rule $f \in F$ and a state $\xi \in E$ with

$$E_x^f \left[\int_0^{\tau_\xi} c(X_t, \pi_t) dt \right] < \infty \tag{5}$$

for all $x \in E$, where $\tau_\xi = \inf\{t \geq 0 \mid X_t = \xi\}$. Then there exist a constant $\bar{\beta} > 0$ and a function $M : E \to I\!R_+$ such that for all $x \in E$ and $0 < \beta < \bar{\beta}$

$$h^\beta(x) = V^\beta(x) - V^\beta(\xi) \leq M(x)$$

and $G_f(x) < \infty$ for all $x \in E$.

Proof: Let $\pi^\beta = (f^\beta, f^\beta, \ldots)$ be the optimal stationary policy for the β-discounted model and denote by (π_t^β) the process of the optimal control, starting in ξ. (π_t) is the process of the control starting in x under policy π. Now define for $t \geq 0$

$$\tilde{\pi}_t^\beta = \begin{cases} \pi_t & \text{, if } t < \tau_\xi \\ \pi_{t-\tau_\xi}^\beta & \text{, if } t \geq \tau_\xi \end{cases}$$

For arbitrary $\beta > 0$ we obtain for $x \in E$

$$V^\beta(x) \leq V_{\tilde{\pi}^\beta}^\beta(x) \leq E_x^f \left[\int_0^{\tau_\xi} c(X_t, \pi_t)\, dt \right] + V^\beta(\xi).$$

Hence we can define $M(x) := E_x^f \left[\int_0^{\tau_\xi} c(X_t, \pi_t)\, dt \right]$ which is finite due to our assumption. From Theorems 4.3, 7.1 in Meyn/Tweedie (1993) we obtain $G_f(x) < \infty$. □

The assumption that $L \leq h^\beta(x)$ for $0 < \beta < \bar{\beta}$, $x \in E$ is clearly fulfilled, if we have monotonicity, i.e. $V^\beta(x) \geq V^\beta(\xi)$ for all $x \in E$ and $0 < \beta < \bar{\beta}$. Another important case where this condition is fulfilled emerges when the cost rate function is coercive (see e.g. Kitaev/Rykov (1995)).

Definition 2:

The cost rate function $c : E \times U \to I\!R_+$ will be called *coercive* when the set $B_r := \{x \in E \mid \inf_{u \in U} c(x, u) \leq r\}$ is compact for all $r \in I\!R_+$.

Remark 1:

Since c is continuous and U compact (Assumption 1), we obtain that $x \mapsto \min_{u \in U} c(x,u)$ is continuous and hence B_r is closed. Therefore, under Assumption 1, a growth condition on c like the one in Assumption 3 is sufficient for the coercivity of c.

Assumption 3:

There exist constants $k \in I\!N$ and $C_1, C_2 \in I\!R_+$ such that for all $z \in Z, u \in U$ and $y \in S$

$$c(y,z,u) \geq C_1 \|y\|^k - C_2.$$

Lemma 7:

Suppose that Assumptions 1 and 2 (i) hold and let $\bar{\beta} > 0$. Assume that there exists an upper semicontinuous function $M : E \to I\!R_+$ such that $-M(x) \leq h^\beta(x) \leq M(x)$ for all $x \in E$ and $0 < \beta < \bar{\beta}$. If the cost rate function satisfies Assumption 3 , then there exists a constant $L \in I\!R$ such that for all $x \in E,\ 0 < \beta < \bar{\beta}$

$$L \leq h^\beta(x).$$

The proof uses ideas of Sennott (1989b) Proposition 3.

Proof: Define $\rho = \limsup_{\beta \downarrow 0} \rho(\beta)$. Note that ρ is finite due to Assumption 2 (i) and Theorem 1. Choose $r > \max\{\rho + \epsilon, \min_u c(\xi, u)\}$ for $\epsilon > 0$. Hence $\xi \in B_r$. Since B_r is compact, M upper semicontinuous and V^β lower semicontinuous due to our assumptions (cf. Bäuerle (1999) Theorem 3) we can define

$$-L = \max_{x \in B_r} M(x), \quad V^\beta(x^\beta) = \min_{x \in B_r} V^\beta(x).$$

From our assumptions we have

$$-M(x) \leq V^\beta(x) - V^\beta(\xi) \leq M(x)$$

for all $0 < \beta < \bar{\beta}$ and $x \in E$. Hence for all $x \in B_r$

$$\beta V^\beta(x) \geq \beta \left(-M(x) + V^\beta(\xi) \right) \geq \beta \left(L + V^\beta(\xi) \right) \tag{6}$$

$$\beta V^\beta(x) \le \beta\left(M(x) + V^\beta(\xi)\right) \le \beta\left(-L + V^\beta(\xi)\right)$$

and $\limsup_{\beta \downarrow 0} \beta(-L + V^\beta(\xi)) = \rho$. Therefore, we can conclude that there exists a $\bar\beta > 0$ such that $\beta V^\beta(x) \le \rho + \epsilon$ for all $x \in B_r$, $0 < \beta < \bar\beta$. In particular $\beta V^\beta(x^\beta) \le \rho + \epsilon$ if $0 < \beta < \bar\beta$. Now suppose $x \notin B_r$ and $0 < \beta < \bar\beta$ and define $\tau := \inf\{t \ge 0 \mid X_t \in B_r\}$ where (X_t) is the state process induced by the β-discounted optimal policy π_t^β. Thus

$$\begin{aligned} V^\beta(x) &\ge E_x^{\pi^\beta}\left[\int_0^\tau e^{-\beta t} c(X_t, \pi_t^\beta)\, dt + e^{-\beta\tau} V^\beta(x^\beta)\right] \\ &\ge E_x^{\pi^\beta}\left[(\rho+\epsilon)\frac{1-e^{-\beta\tau}}{\beta} + e^{-\beta\tau} V^\beta(x^\beta)\right] \ge V^\beta(x^\beta). \end{aligned}$$

Notice, that the statement is true even if $\tau = \infty$. Hence we have for $x \in B_r$ from (6) that $V^\beta(x) - V^\beta(\xi) \ge L$ and for $x \notin B_r$

$$V^\beta(x) - V^\beta(\xi) \ge V^\beta(x^\beta) - V^\beta(\xi) \ge L$$

which implies the statement. □

5 Applications

In this section we apply our results to a multi-product manufacturing system and multi-class queueing networks.

A) Manufacturing systems

The example is taken from Sethi et al. (1998) (cf. also Sethi/Zhang (1994)). We have a number of parallel machines for manufacturing which are subject to random breakdown and repair. Each machine is capable of producing any of K different products. The vector $y = (y_1, \ldots, y_K)$ gives the inventory/backlog of each product and we assume $S = \mathbb{R}^K$. $\lambda(z) \in \mathbb{R}_+, z \in Z$ gives the production capacity of the system that is available. The vector $u \in U = \{u \in [0,1]^K \mid \sum_{j=1}^K u_j \le 1\}$ gives the percentages of the production capacity that are assigned to each of the products. If we denote by $\mu \in \mathbb{R}_+^K$ the constant demand rate, the dynamics of the system are given by

$$b^z(u) = \lambda(z)u - \mu.$$

The function $c : I\!R^{2K} \to I\!R_+$ denotes the surplus (inventory/backlog) and production cost. In order to apply our results, we have to impose the following assumption on the cost rate function

Assumption 4:

(i) $(y,u) \mapsto c(y,z,u)$ is convex for all $z \in Z$.

(ii) there exist constants $k \in I\!N$ and $C_0 \in I\!R_+$ such that for all $z \in Z$, $u, u' \in U$ and $y, y' \in S$

$$|c(y,z,u) - c(y',z,u')| \leq C_0(1 + \|y\|^k + \|y'\|^k)(\|y - y'\| + \|u - u'\|).$$

(iii) there exist constants $l \in I\!N$ and $C_1, C_2 \in I\!R_+$ such that for all $z \in Z$, $u \in U$ and $y \in S$

$$c(y,z,u) \geq C_1\|y\|^l - C_2.$$

Moreover, we need the following stability condition

Assumption 5:

Suppose that ν is the stationary distribution of the environment process (Z_t), i.e. $\nu \geq 0$ satisfies $\nu Q = 0$, $\sum_z \nu_z = 1$. Then we assume $\sum_z \lambda(z)\nu_z > \sum_{j=1}^K \mu_j$.

It is easy to see that in this model our Assumption 1 is fulfilled. Since the state trajectory can grow at most linear and the cost rate function is bounded by a polynom (see Assumption 4 (iii)), the average cost as well as $V_\pi^\beta(x)$ are finite for all policies. In Sethi et al. (1998) Theorem 3 (cf. also Sethi et al. (1997) Theorem 3.3) it has been shown that with $\xi = (0,0) \in E$ (w.l.o.g. suppose $0 \in Z$) we have

$$|h^\beta(x)| \leq C_0(1 + \|y\|^{k+2}) =: M(y)$$

for all $x \in E, \beta > 0$, where $C_0 \in I\!R_+$ is independent of β. Since the assumptions of Remark 1 are fulfilled, we obtain with Lemma 7 that $h^\beta(x) \geq L$. Hence altogether

Assumption 2 is fulfilled and Theorems 4 and 5 are valid.

Let us finally look at the special case of a one-product system, i.e. we have $K = 1$. In this case it is possible to show that a threshold policy is optimal (see e.g. Sethi et al. (1997)). In addition we show that the optimal threshold is a limit of thresholds which are optimal in the discounted models as the discount factor approaches zero. We need one further assumption:

Assumption 6:

The cost rate function is of the form $c(y, z, u) = c_1(y) + \hat{c}u$ with $\hat{c} \in \mathbb{R}_+^K$.

For the β-discounted problem it is possible to show (cf. Sethi et al. (1997)) that the optimal policy is given by a threshold feedback control $\varphi^\beta : E \to U$, i.e. there exists a function $S^\beta : Z \to \mathbb{R}$ such that

$$\varphi^\beta(x) = \begin{cases} 1, & y < S^\beta(z) \\ \min\{1, \frac{\mu}{\lambda(z)}\}, & y = S^\beta(z) \\ 0, & y > S^\beta(z). \end{cases}$$

If $|Z| = 2$, the function $S^\beta(z)$ can be computed explicitely. This has been done in Akella/Kumar (1986). For arbitrary Z it is possible to derive monotonicity properties of $S^\beta(z)$. In Sethi/Zhang (1994) one finds statements if (Z_t) is a birth-and-death process. For a more general concept using stochastic orderings, see Rajagopal et al. (1995). In the average cost case we obtain now

Corollary 8:

Assume further that $K = 1$ and that assumptions 4, 5 and 6 hold. Then the optimal policy in the c-average cost model is given by a threshold feedback control φ, i.e. there exists a function $S : Z \to \mathbb{R}$ such that

$$\varphi(x) = \begin{cases} 1, & y < S(z) \\ \min\{1, \frac{\mu}{\lambda(z)}\}, & y = S(z) \\ 0, & y > S(z). \end{cases}$$

Moreover, there exists a sequence $\beta_m \to 0$ such that $S^{\beta_m}(z) \to S(z)$ for $z \in Z$, where S^{β_m} is the optimal threshold function in the β_m-discounted model.

Proof: We can choose in the proof of Theorem 4 a subsequence $\{\beta_{n_m}\}$ of $\{\beta_n\}$ such that $S^{\beta_{n_m}}(z) \to S(z)$ for all $z \in Z$ and $m \to \infty$ (cf. also Schäl (1993) section 4). Next we have to verify that $f^{\beta_{n_m}}(x) \to f(x)$ for $m \to \infty$, where f is constructed from a feedback control φ with threshold function S. Theorem 5 then implies the statement. For the convergence result $f^{\beta_{n_m}}(x)$ is interpreted as an element in $\mathcal{R} := \{r : \mathbb{R}_+ \to \mathbb{P}(U) \mid r \text{ measurable}\}$, where $\mathbb{P}(U)$ is the set of all probability measures on U. To prove convergence $r_n \to r$, for $r_n, r \in \mathcal{R}$ we have to show that

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

for all measurable functions $\psi : \mathbb{R}_+ \times U \to \mathbb{R}$ such that $u \mapsto \psi(t,u)$ is continuous for all $t \geq 0$ and $\int_0^\infty \sup_{u \in U} |\psi(t,u)|dt < \infty$. In our case this makes it necessary to distinguish between several cases. We will only look at the case $\lambda(z) > \mu$ and $y > S(z) \in \mathbb{R}$. We have to show

$$\int_0^\infty \int_U \psi(t,u) f^{\beta_{n_m}}(x)(t,du)dt \to \int_0^\infty \int_U \psi(t,u) f(x)(t,du)dt$$

for all measurable functions ψ with the preceding properties. W.l.o.g. suppose $y - S(z) > 2\epsilon$ for $\epsilon > 0$. Choose $N_0(\epsilon)$ big enough such that for all $m \geq N_0(\epsilon)$:

$$|S^{\beta_{n_m}}(z) - S(z)| \leq \epsilon$$

and thus $y > S^{\beta_{n_m}}(z)$ for all $m \geq N_0(\epsilon)$. Hence we obtain with $t_m(x) := (y - S^{\beta_{n_m}}(z))/\mu$

$$\left| \int_0^{t(x)} \psi(t,0)dt + \int_{t(x)}^\infty \psi(t, \frac{\mu}{\lambda(z)})dt - \int_0^{t_m(x)} \psi(t,0)dt - \int_{t_m(x)}^\infty \psi(t, \frac{\mu}{\lambda(z)})dt \right| \leq$$

$$\leq \int_{t(x)}^{t_m(x)} |\psi(t,0)|dt + \int_{t(x)}^{t_m(x)} |\psi(t, \frac{\mu}{\lambda(z)})|dt \to 0 \text{ for } m \to \infty,$$

since $t_m(x) \to t(x)$ for $m \to \infty$ which implies the statement. $\square$

B) Stochastic multiclass fluid networks

A special case of SFPs are stochastic multiclass fluid networks (see e.g. Dai (1995)). They consist of J service stations, each with a single server and $K \geq J$ fluid classes. $C(j) \subset \{1, \ldots, K\}$ are the fluid classes which have to be processed at station j. The external inflow rate of class k at time t is given by $\alpha_k(Z_t)$, where (Z_t) is our environment process. We suppose that (Z_t) is such that $(\alpha_k(Z_t))$ is an irreducible continuous-time Markov chain itself and the processes $(\alpha_1(Z_t)), \ldots, (\alpha_K(Z_t))$ are stochastically independent of each other. This is e.g. fulfilled, if $(Z_t) = (Z_1(t), \ldots, Z_K(t))$, where $(Z_1(t)), \ldots, (Z_K(t))$ are independent and $\alpha_k(Z_t) = \alpha_k(Z_k(t))$. The state process $(Y_t) = (Y_1(t), \ldots, Y_K(t))$ describes the buffer contents of the different classes over time. We suppose that $S = \mathbb{R}_+^K$. The set U consists now of all possible server allocations to the classes, i.e. $U := \{u \in [0,1]^K \mid \sum_{k \in C(j)} u_k \leq 1, \ j = 1, \ldots, J\}$. For $u \in U$, u_k gives the fraction of the responsible server which is assigned to class k. The potential service rate of class k is $\mu_k > 0$. Thus, if the server allocation $u \in U$ is chosen, the outflow rate of class k is $\mu_k u_k$. A fraction p_{ki} of the fluid which is leaving class k is routed to class i. Therefore, $\sum_{k=1}^K p_{ki} \mu_k u_k$ is the internal inflow rate of class i. Throughout we will suppose that the routing matrix $P = (p_{ki})$ is transient, i.e. $P^n \to 0$ for $n \to \infty$. This implies in particular that $(I - P)^{-1} = \sum_{n=0}^\infty P^n \geq 0$, where I is the identity matrix. If we denote $A := diag(\mu)(I - P)$, the drift of the network is given by

$$b^z(u) = \alpha(z) - uA.$$

For the cost rate function we take linear holding cost, i.e. $c(x, u) = \sum_{k=1}^K c_k y_k$ with $c_k \geq 0$. The optimization problem is now to find a server allocation such that the average holding cost in the system are minimized. We assume that the network is such that $V^\beta(y, z)$ is increasing in y. We will show that Assumptions 1 and 2 of Section 3 are satisfied for this model. Assumption 1 is obviously fulfilled. It is easy to see that every policy π satisfies $V_\pi^\beta(x) < \infty$ for all $x \in E$, $\beta > 0$. As far as Assumption 2 is concerned, we need a further stability condition. Suppose that $\lambda(z) = (\lambda_1(z), \ldots, \lambda_K(z))$ is the nominal total arrival rate to the different classes, i.e. $\lambda(z)$ is the solution of the equation $\lambda_i(z) = \alpha_i(z) + \sum_{k=1}^K \lambda_k(z) p_{ki}$. In matrix notation this gives $\lambda(z) = \alpha(z) + \lambda(z)P$. Since

P is transient we obtain

$$\lambda(z) = \alpha(z)(I - P)^{-1}.$$

The traffic intensity $\rho_j(z)$ at station j is then defined by

$$\rho_j(z) = \sum_{k \in C(j)} \frac{\lambda_k(z)}{\mu_k}.$$

We will now assume that the usual traffic conditions are satisfied on average, i.e.

Assumption 7:

Suppose that ν is the stationary distribution of the environment process (Z_t). Then we assume $\sum_{z \in Z} \rho_j(z)\nu_z < 1$ for $j = 1, \dots, J$.

W.l.o.g. we assume that $0 \in Z$ and $\alpha_k(0) < \sum_{z \in Z} \nu_z \alpha_k(z),\ k = 1, \dots, K$. Hence $(0,0) \in E$ is the state, where all buffers are empty and the environment process is in state 0.

Lemma 9:

Suppose that Assumption 7 is valid. Then there exists a decision rule $f \in F$ such that

$$E_x^f \left[\int_0^{\tau_{(0,0)}} cY_t dt \right] \leq C(1 + \|y\|^2)$$

for all $x \in E$, where C is independent of y.

Proof: Since Y_t can grow at most linear, there exists a constant $\tilde{c} \in \mathbb{R}_+$ such that

$$\begin{aligned} E_x^f \left[\int_0^{\tau_{(0,0)}} cY_t \, dt \right] &\leq E_x^f \left[\int_0^{\tau_{(0,0)}} cy + \tilde{c}t \, dt \right] \\ &= cyE_x^f[\tau_{(0,0)}] + \frac{1}{2}\tilde{c}E_x^f[\tau_{(0,0)}^2]. \end{aligned}$$

Thus, we have to show that $E_x^f[\tau_{(0,0)}] \leq \tilde{C}(1 + \|y\|^2)$ and $E_x^f[\tau_{(0,0)}^2] \leq \tilde{C}(1 + \|y\|^2)$. The first inequality has been shown in Bäuerle (2000) Theorem 6. For the second inequality we proceed in the same way. The proof contains ideas of Sethi et al. (1997), Lemma 3.1. From Bäuerle (2000) we know that under the stability assumption, there exists a decision rule $f \in F$ such that $E_x^f[\tau_1^2] \leq C_0(1 + \|y\|^2)$ for all $x \in E$, where $\tau_1 = \inf\{t \geq 0 \mid Y_t = 0\}$.

Let $\mathcal{Z} := \{z \in Z \mid \alpha_k(z) \le \sum_{z' \in Z} \nu_{z'} \alpha_k(z'),\ k = 1, \dots, K\}$. Note that $\mathcal{Z} \neq \emptyset$ since $0 \in \mathcal{Z}$. Let us now define the following sequence of stopping times:

$$\sigma_1 := \inf\{t \ge \tau_1 \mid Z_t \notin \mathcal{Z}\}$$

$$\tau_n := \inf\{t > \sigma_n \mid Y(t) = 0\}$$

$$\sigma_n := \inf\{t \ge \tau_n \mid Z_t \notin \mathcal{Z}\}.$$

Then $0 \le \tau_1 \le \sigma_1 \le \ldots < \tau_n \le \sigma_n$ and $Y_t = 0$ for $t \in [\tau_n, \sigma_n),\ n \in I\!N$. Moreover, since (Z_t) is positive recurrent

(i) $P_x^f(Z_s \neq 0,\ \tau_n \le s < \sigma_n) \le \delta < 1$, for all $n \in I\!N$.

(ii) $E_x^f[(\sigma_n - \tau_n)^2] \le C_1$, for all $n \in I\!N,\ x \in E$.

From (i) we conclude that

$$\begin{aligned} P_x^f(\tau_{(0,0)} > \sigma_n) &= P_x^f(\tau_{(0,0)} > \sigma_n, \dots, \tau_{(0,0)} > \sigma_1) = P_x^f(\tau_{(0,0)} > \sigma_n \mid \tau_{(0,0)} > \sigma_{n-1}) \\ &\quad \cdot \ldots \cdot P_x^f(\tau_{(0,0)} > \sigma_2 \mid \tau_{(0,0)} > \sigma_1) P_x^f(\tau_{(0,0)} > \sigma_1) \le \delta^n. \end{aligned}$$

From (ii), the Cauchy Schwarz inequality and $\max_z E^f_{(0,z)}[(\tau_n - \sigma_{n-1})^2] \le C_2$ we get

$$E_x^f[\tau_n^2] = E_x^f[(\tau_1 + (\sigma_1 - \tau_1) + (\tau_2 - \sigma_1) + \ldots + (\tau_n - \sigma_{n-1}))^2] \le n^2 C_3 (1 + \|y\|^2)$$

where C_3 is independent of x and n. Altogether we obtain now

$$\begin{aligned} E_x^f[\tau_{(0,0)}^2] &= 2\int_0^\infty t P_x^f(\tau_{(0,0)} > t)\,dt = 2\sum_{n=1}^\infty E_x^f\left[\int_{\tau_{n-1}}^{\tau_n} t P_x^f(\tau_{(0,0)} > t)\,dt\right] \le \\ &\le E_x^f[\tau_1^2] + 2\sum_{n=2}^\infty E_x^f\left[\int_{\tau_{n-1}}^{\tau_n} t P_x^f(\tau_{(0,0)} > \sigma_{n-2})\,dt\right] \le \\ &\le C_0(1 + \|y\|^2) + \sum_{n=2}^\infty \delta^{n-2} E_x^f\left[\tau_n^2 - \tau_{n-1}^2\right] \le \\ &\le C_0(1 + \|y\|^2) + \sum_{n=2}^\infty C_4\, n^2 \delta^{n-2}(1 + \|y\|^2) \le C_5(1 + \|y\|^2) \end{aligned}$$

and the assertion follows. □

Thus, according to Lemma 7 we get the upper bound for $h^\beta(x) = V^\beta(x) - V^\beta(0,0)$ and that $G_f(x) < \infty$. The lower bound follows since the discounted value functions $V^\beta(y,z)$ are increasing in y. In particular $V^\beta(y,z) \ge V^\beta(0,z)$.

Altogether we have shown that the assumptions of Section 3 are valid under the stability Assumption 7 for this model. Hence, there exists in particular an average optimal decision rule. Moreover, it has been shown in Bäuerle/Rieder (2000) and Bäuerle/Stidham (2000) that for the single-server model, i.e. $J = 1$ a priority index rule minimizes the β-discounted cost. More precisely, it is possible to compute indices I_k, $k = 1, \ldots, K$ for each class such that it is optimal to drain the buffers according to the priority given by the indices, from highest to smallest. The indices I_k are independent of the interest rate β. Thus, we obtain with Theorem 4 c) that the index rule is also optimal for the average cost problem.

References

Akella RA and Kumar PR (1986) Optimal control of production rate in a failure prone manufacturing system. *IEEE Trans. Automa. Control AC* **31** 116-126.

Bäuerle N (1999) Discounted stochastic fluid programs. To appear in *Mathem. Operations Res.*

Bäuerle N (2000) On positive Harris recurrence of stochastic fluid networks. *Preprint.*

Bäuerle N and Rieder U (2000) Optimal control of single-server fluid networks. To appear in *Queueing Systems.*

Bäuerle N and Stidham S (2000) Sample-path conservation laws for single-server fluid networks. *Preprint.*

Bilecki T and Kumar PR (1988) Optimality of zero-inventory policies for unreliable manufacturing systems. *Operations Research* **36** 532-546.

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

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

Hernández-Lerma O and Lasserre JB (1996) *Discrete-time Markov control processes.* Springer-Verlag, New York.

Hordijk A and Van der Duyn Schouten FA (1983) Average optimal policies in Markov decision drift processes with applications to a queueing and a replacement model.

Adv. Appl. Probab. **15** 274-303.

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

Meyn SP and Tweedie RL (1993) Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **25** 518-548.

Presman E, Sethi S and Zhang Q (1995) Optimal feedback production planning in a stochastic N-machine flowshop. *Automatica* **31** 1325-1332.

Rajagopal S (1995) *Optimal control of stochastic fluid-flow systems with applications to telecommunication and manufacturing systems.* PhD Dissertation at the University of North Carolina, Chapel Hill.

Rajagopal S, Kulkarni VG and Stidham S (1995) Optimal flow control of stochastic fluid-flow systems. *IEEE Journal on selected areas in Communications* **13** 1219-1228.

Rockafellar RT (1972) *Convex analysis.* Princeton University Press, Princton.

Rolski T, Schmidli H, Schmidt V and J Teugels (1999) *Stochastic processes for insurance and finance.* John Wiley & Sons. Chichester.

Schäl M (1993) Average optimality in dynamic programming with general state space. *Math. Operations Res.***18** 163-172.

Sennott LI (1989a) Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Res.* **37** 626-633.

Sennott LI (1989b) Average cost semi-Markov decision processes and the control of queueing systems. *Probab. Engineering and Inform. Sci.* **3** 247-272.

Sethi SP, Suo W, Taksar MI and Yan H (1998) Optimal production planning in a multi-product stochastic manufacturing system with long-run average cost. *Discrete Event Dynamic Systems: Theory and Applications* **8** 1-18.

Sethi SP, Suo W, Taksar MI and Zhang Q (1997) Optimal production planning in a stochastic manufacturing system with long-run average cost. *J. Optim. Theory Appl.* **92** 161-188.

Sethi SP and Zhang Q (1994) *Hierarchical decision making in stochastic manufacturing systems.* Birkhäuser, Boston.

Veatch MH and Caramanis MC (1999) Optimal manufacturing flow controllers: zero-inventory policies and control switching sets. *IEEE Transactions on Automatic Control* **44** 779-783.