Optimal Control of single-server fluid networks

Nicole Bäuerle [a] Ulrich Rieder [b]

[a] *Department of Mathematics VII, University of Ulm, D-89069 Ulm, Germany*
E-mail: baeuerle@mathematik.uni-ulm.de
[b] *Department of Mathematics VII, University of Ulm, D-89069 Ulm, Germany*
E-mail: rieder@mathematik.uni-ulm.de

We consider a stochastic single-server fluid network with both a discounted reward and a cost structure. It can be shown that the optimal policy is a priority index policy. The indices coincide with the optimal indices in a Semi-Markovian Klimov problem. Several special cases like single-server re-entrant fluid lines are considered. The approach we use is based on sample path arguments and Pontryagins maximum principle.
Keywords: stochastic fluid model, index policy, Klimov index, largest remaining index algorithm, maximum principle, sample path argument, re-entrant fluid lines

1. Introduction

The model under consideration is a stochastic single-server fluid network. It consists of a finite number of buffers and a single server who has to be allocated to the buffers. The input flow can come from both external sources as well as internal transitions which occur according to a routing matrix. The external arrival process of fluid is driven by a continuous-time Markov chain with finite state space. We consider both a discounted reward and a discounted cost model.

The motivation to study fluid models is twofold. First, in many applications which involve hierarchical decision making, fluid models seem to be natural, since the frequency of occurence of different types of events is different. Therefore, quantities that vary much faster than others are modeled in a deterministic way by replacing them with their averages. This is a common technique in manufacturing systems (see Sethi/Zhang [8]). Second, fluid models capture the asymptotic behavior of discrete stochastic queueing systems. In Chen/Mandelbaum [3] it is shown that a class of queueing networks converges under appropriate time and space scaling to fluid networks. Moreover, recent results have shown a close connection between stability of stochastic networks and stability of their fluid models. This implies also the hope that an optimal control for a fluid model - if translated - can be used as a good heuristic for the discrete stochastic model (see also Meyn [6]).

Indeed, with our approach we can show that the optimal control in the discounted single-server fluid network is a priority index policy. The indices are independent of the discount factor and coincide with the optimal indices in a Semi-Markovian Klimov problem under the average reward criterion. Under additional assumptions this has already been discovered for the deterministic fluid model in Chen/Yao [4]. Using a linear programming approach the authors there showed that the index policy is a myopic solution of the optimization problem and gave conditions under which the myopic solution is also globally optimal. In contrast, we use a sample path argument and Pontryagins maximum principle to establish the optimality even in a stochastic fluid setting. Moreover, we can show that the conditions needed in Chen/Yao [4] are always fulfiled under reasonable assumptions. Note that recent studies in this spirit also include Avram/Bertsimas/Ricard [2], Avram [1] and Weiss [12].

The paper is organized as follows: in section 2 we give a precise mathematical formulation of the optimization problem, where we introduce the cost and the reward model. As in the well-known discrete-time stochastic single-server problem we show that there is a relation between these two models. In addition we prove that the value functions are convex and concave respectively. In section 3 we define the indices recursively by a so-called largest remaining index algorithm. From this representation we can see that the indices coincide with the indices of an adequately defined Klimov problem. This enables us to derive easily some bounds and properties of the indices like monotonicity. The next section contains the definition of the index policy and states the optimality of it under some mild and natural assumptions like stability. Moreover, we consider the following special cases of the model: zero-routing which implies that the index policy is the well-known μc-rule, a deteriorating case where the myopic policy is optimal and a single-server re-entrant line which has also been considered in Weiss [12]. The proof of the optimality of the priority index policy is given in section 5.

2. The Model

Consider the following single-server fluid network: there are N queues with infinite buffers. Each buffer receives a continuous stochastic inflow in the following way: Let $\{Z_t,\ t \geq 0\}$ be a continuous time Markov chain with finite state space Z and generator $Q = (q_{zz'})$. As usual define $q_z = -q_{zz}$. We will interpret Z_t as the state of the environment at time t and we will call $\{Z_t,\ t \geq 0\}$ the environment process. Denote by $T_1, T_2, \ldots$ the random jump times of $\{Z_t,\ t \geq 0\}$. Given, the environment process at time t is in state a, there is a deterministic inflow into buffer j at rate $\lambda_j^z \geq 0,\ j = 1, \ldots, N$. Let us define $\lambda^z = (\lambda_1^z, \ldots, \lambda_N^z)$. A single server has to be splitted among the buffers. The potential service rate of buffer j is assumed to be $\mu_j > 0,\ j = 1, \ldots, N$ which means that if a fraction $u_j \in (0,1)$ of the server is allocated to buffer j, there is an output rate of $u_j \mu_j$ if the buffer content is greater zero. If it is zero, the actual output rate is equal to the minimum of $u_j \mu_j$ and the input rate. For abbreviation denote the matrix $D = diag(\mu_j)$ as the diagonal matrix with elements μ_j on the diagonal. The fluid that is leaving buffer j is divided and a fraction of $p_{ji} \in [0,1),\ i = 1, \ldots, N$ is instantaneously flowing into buffer i. Throughout this paper we assume that

(A 1) $\sum_{i=1}^N p_{ji} < 1, \quad \text{for all } j = 1, \ldots, N.$

i.e. a positive fraction of $1 - \sum_{i=1}^N p_{ji}$ is leaving the system. Denote the matrix $P = (p_{ji})$ and define $\mathcal{U} = \{u \in [0,1]^N \mid \sum_{j=1}^N u_j \leq 1\}$. Hence, given a fixed server allocation $u \in \mathcal{U}$ and a fixed environment state a, the **input rate** into buffer $j,\ j = 1, \ldots, N$ is $\lambda_j^z + \sum_{i=1}^N p_{ij} \mu_i u_i$ and the **output rate** is equal to $\mu_j u_j$. In matrix notation this is $\lambda^z + P^T D u$ and Du respectively. The **state of the system** at time t is $(Y_t, Z_t) = (Y_1(t), \ldots, Y_N(t), Z_t)$ with the interpretation that $Y_j(t)$ is the buffer content of buffer j at time t, $1 \leq j \leq N$. Hence $\mathbb{R}_+^N \times Z$ is the state space of the system. If the state of the system is (y, z), we define

$$\mathcal{U}(y,z) = \{u \in \mathcal{U} \mid \lambda_j^z - \mu_j u_j + \sum_{i=1}^N p_{ij} \mu_i u_i \geq 0, \text{ whenever } y_j = 0\}. \tag{1}$$

For an arbitrary function $u \in \mathcal{M} = \{u : \mathbb{R}_+ \to \mathcal{U} \mid u \text{ measurable }\}$ denote $\phi(y,z,u)(t) = y + \int_0^t \lambda^z - Du_s + P^T Du_s\, ds$. If we introduce the matrix $A = D(I - P)$, where I is the identity matrix then we can simply write

$$\phi(y,z,u)(t) = y + \int_0^t \lambda^z - A^T u_s\, ds \tag{2}$$

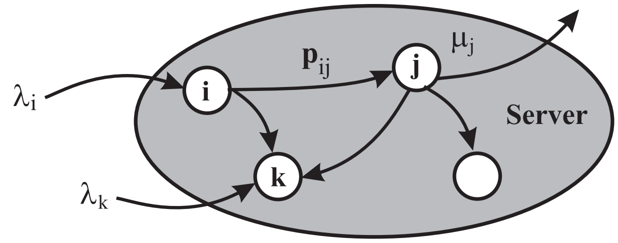

Figure 1. Single-server network

Hence $D(y,z) = \{u \in \mathcal{M} \mid u_t \in \mathcal{U}(\phi(y,z,u)(t),z)\}$ is defined as the **set of admissible controls**. At each jump time of the environment process we have to choose an admissible control. This is done by a **decision rule** from the set $F = \{f : I\!R_+^N \times Z \to \mathcal{M} \mid f$ measurable, $f(y,z) \in D(y,z)$ for all $(y,z) \in I\!R_+^N \times Z\}$. For $f \in F$ let $\phi^f(y,z)(t) = \phi(y,z,f(y,z))(t)$. A **policy** π is now defined by a sequence of decision rules i.e. $\pi = (f_n)$, where $f_n \in F$ for all $n \in I\!N$. If we define by

$$\pi_t = f_n(Y_{T_n}, Z_{T_n})(t - T_n) \quad \text{for } T_n \le t < T_{n+1}, \tag{3}$$

then π_t is the control we have to apply at time t. Analogously we define

$$Y_t = \phi^{f_n}(Y_{T_n}, Z_{T_n})(t - T_n) \quad \text{for } T_n \le t < T_{n+1}. \tag{4}$$

We are now interested in two different optimality criterions:

In the **reward model** we suppose that we obtain a reward of $r_j \in I\!R$ for each unit of fluid from buffer j that is processed. Denote $r = (r_1, \ldots, r_N)$. The aim is to maximize the expected discounted reward of the system over an infinite horizon. Hence, for a given admissible policy π we define for $\beta > 0$

$$V_\pi(y,z) = E_{y,z}\left[\int_0^\infty e^{-\beta t} r\pi_t \, dt\right]$$

where $Y_0 = y,\ Z_0 = z$ and π_t is given by (3).

In the **cost model** we suppose that a linear cost of rate $c_j \in I\!R$ is incurred, when holding fluid in buffer j. Denote $c = (c_1, \ldots, c_N)$. We are interested in finding a control which minimizes the expected discounted cost of the system over an infinite horizon. Hence, for a given admissible policy π we define for $\beta > 0$

$$C_\pi(y,z) = E_{y,z}\left[\int_0^\infty e^{-\beta t} cY_t \, dt\right]$$

where Y_t is given by (4), $Y_0 = y,\ Z_0 = z$. Hence the optimization problems are

$$V(y,z) = \sup_{\pi \in F^\infty} V_\pi(y,z)$$

$$C(y,z) = \inf_{\pi \in F^\infty} C_\pi(y,z)$$

An important special case is obtained, when λ^z is equal to λ for all $a \in Z$. Then the inflow process is constant and the model is purely deterministic. In this case the reward model is the following control problem

$$\begin{cases} \int\limits_0^\infty e^{-\beta t} r u_t \, dt \to \max \\ \dot{y}_t = \lambda - A^T u_t \\ y_t \geq 0 \\ u_t \in \mathcal{U} \end{cases}$$

As in the discrete time stochastic model there is a certain connection between the cost and the reward model (see e.g. Weishaupt [11], Tcha/Pliska [9]).

If we define $r_j = \mu_j[c_j - \sum_{i=1}^N p_{ji} c_i]$ or in matrix notation $r = Ac$, we obtain

Lemma 1. Let π be a policy and $y \geq 0$ the initial inventory of the fluid network. Then we obtain for $a \in Z$

$$\beta C_\pi(y,z) = cy + E\left[\int_0^\infty e^{-\beta s} c \lambda^{Z_s} \, ds\right] - V_\pi(y,z)$$

In particular, in the deterministic setting, if all λ^z are equal λ, then

$$\beta C_\pi(y) = cy + \frac{1}{\beta} c\lambda - V_\pi(y)$$

Proof. The proof is using a sample path argument. Let ω be fixed. Denote by $Z(t,\omega)$ the path of the environment process and by $Y(t,\omega)$ the path of the buffer contents under policy π, where $Y(0,\omega) = y,\ Z(0,\omega) = z$. By definition we have

$$\begin{aligned} C_\pi(y,z)(\omega) &= \int_0^\infty e^{-\beta t} \sum_{j=1}^N c_j Y_j(t,\omega) \, dt \\ &= \int_0^\infty e^{-\beta t} \sum_{j=1}^N c_j \left(\int_0^t \dot{Y}_j(s,\omega) \, ds + y_j\right) \, dt \\ &= \frac{1}{\beta} \int_0^\infty e^{-\beta s} \sum_{j=1}^N c_j \dot{Y}_j(s,\omega) \, ds + \frac{1}{\beta} \sum_{j=1}^n c_j y_j \end{aligned}$$

where the second equality follows from changing the order of integration. Substituting $\dot{Y}_j(s,\omega)$ by the expression $\lambda_j^{Z(s,\omega)} - \mu_j\pi_j(s,\omega) + \sum_{i=1}^N p_{ij}\mu_i\pi_i(s,\omega)$ we obtain

$$\int_0^\infty e^{-\beta s}\sum_{j=1}^N c_j\dot{Y}_j(s,\omega)\,ds = \int_0^\infty e^{-\beta s}\sum_{j=1}^N \pi_j(s,\omega)\mu_j\left(-c_j+\sum_{i=1}^N p_{ji}c_i\right)\,ds + \int_0^\infty e^{-\beta s}\sum_{j=1}^N c_j\lambda_j^{Z(s,\omega)}\,ds =$$

$$= -V_\pi(y,z)(\omega) + \int_0^\infty e^{-\beta s}\sum_{j=1}^N c_j\lambda_j^{Z(s,\omega)}\,ds$$

where we have used the definition of r_j for the last equality. Taking expectation, we obtain the desired result. □

A further interesting property of the value functions is the following:

Lemma 2. $C(\cdot,z)$ is convex and $V(\cdot,z)$ is concave for all $z\in Z$.

Proof. We restrict on showing the assertion on C, the rest follows with Lemma 1.

Let $y,y'\geq 0$ and $\alpha\in[0,1]$. Fix ω and suppose that π and π' are the paths of the optimal policies for start in y and y' respectively. Define $\hat{\pi}_t = \alpha\pi_t + (1-\alpha)\pi'_t,\ t\geq 0$. Hence it holds that $\hat{\pi}_t\in\mathcal{U},\ t\geq 0$. Use $\hat{\pi}$ as a control for start in $\hat{y}=\alpha y+(1-\alpha)y'$ and denote by $\hat{Y}(t,\omega),(\hat{Y}(0,\omega)=\hat{y})$ the associate trajectory, analogously for $Y(t,\omega),(Y(0,\omega)=y)$ and $Y'(t,\omega),(Y'(0,\omega)=y')$. Hence

$$\hat{Y}(t,\omega) = \hat{y} + \int_0^t \lambda^{Z(s,\omega)} - D\hat{\pi}(s,\omega) + P^T D\hat{\pi}(s,\omega)\,ds = \alpha Y(t,\omega) + (1-\alpha)Y'(t,\omega)$$

and in particular $\hat{Y}(t,\omega)\geq 0$ for all $t\geq 0$, since π and π' are admissible. Moreover,

$$C_{\hat{\pi}}(\hat{y},z)(\omega) = \int_0^\infty e^{-\beta t}c\hat{Y}(t,\omega)\,dt = \alpha C(y,z)(\omega) + (1-\alpha)C(y',z)(\omega)$$

and taking expectation, we obtain

$$C(\alpha y+(1-\alpha)y',z) = C(\hat{y},z) \leq C_{\hat{\pi}}(\hat{y},z) = \alpha C(y,z) + (1-\alpha)C(y',z)$$

which is the assertion. □

3. Definition and Properties of the Indices

Our aim is to prove that the optimal policy is a priority index policy. In this section we give a definition of the indices and show several important properties.

Due to assumption (A1) we have that $(I-P)^{-1} = \sum_{n=0}^{\infty} P^n \geq 0$ and hence

$$A^{-1} = \sum_{n=0}^{\infty} P^n D^{-1} \geq 0 \tag{5}$$

For means of short notation, let us introduce the following abbreviation: For a subset $S \subset \{1, \ldots, N\}$ we denote

$$a_i^S = (-a_{ij})_{j\in S} = (\mu_i p_{ij})_{j\in S},\ i \notin S \quad \text{and} \quad A_S = (a_{ij})_{i,j\in S}.$$

An analogous definition is used for vectors. Obviously the relation in (5) holds for arbitrary submatrices A_S.

Now we will give a recursive definition of the indices $I^1, \ldots, I^N$, the so-called **largest remaining index algorithm** (the name will be justified by Theorem 4 a)). By e we denote the vector consisting of 1's only - the dimension should be clear from the context.

Algorithm 3 Largest remaining index algorithm:.

(i) $I^1 = \max_{1\leq j\leq N}\ r_j,\ \ i_1 = \mathbf{argmax}_{1\leq j\leq N}\ r_j,\ \ S_1 = \{i_1\}$.

(ii) For $k = 1, \ldots, N-1$ **let**

$$I_j^{k+1} = \frac{r_j + a_j^{S_k} A_{S_k}^{-1} r_{S_k}}{1 + a_j^{S_k} A_{S_k}^{-1} e_{S_k}}, \quad j \notin S_k$$

$$I^{k+1} = \max_{j\notin S_k} I_j^{k+1}, \quad i_{k+1} = \mathbf{argmax}_{j\notin S_k}\ I_j^{k+1}$$

Set $S_{k+1} = S_k + \{i_{k+1}\}$.

Buffer i_k is now assigned the index $I(i_k) = I^k,\ k = 1, \ldots, N$. The indices have the following nice interpretations:

In the fluid setting: I^1 is simply the maximal reward rate in the model. Suppose that the indices $I^1, \ldots, I^k$ have already been determined. Given that we have to keep the buffer contents of the buffers in S_k at zero we can now look at the reduced network which consists of the buffers in $\{1,\dots,N\} - S_k$. If we allocate a unit of the server to buffer $j \notin S_k$, in order to keep the buffers in S_k empty we have to assign to them a server capacity u_{S_k} which can be computed from

$$0 = A_{S_k}^T u_{S_k} - a_j^{S_k}.$$

Therefore $u_{S_k} = a_j^{S_k} A_{S_k}^{-1}$. Hence I_j^{k+1} is the reward rate of buffer j in the reduced network.

In a Semi-Markovian setting: consider the following Semi-Markovian single-server network (cf. Walrand [10]). Arrivals at node j are according to a Poisson process at rate λ_j. Service times are independent in all queues and have mean $\frac{1}{\mu_j}$ in queue j. When a customer is served in queue j a reward r_j is received and the customer is sent to queue i with probability p_{ji}, $i = 1,\dots,N$ and leaves the network otherwise. The objective is to maximize the average reward of the system over all nonpreemptive service policies. This problem gives exactly the same indices. Moreover, let us look at a particular customer and denote by $X(t)$ the location of him at time t. Define the stopping time $\tau_k = \inf\{t \geq 0 \mid X(t) \notin S_k\}$ as his exit time from the set S_k, then

$$E_j\left[\int_0^{\tau_k} r_{X(t)}\,dt\right] = r_j + a_j^{S_k} A_{S_k}^{-1} r_{S_k}.$$

and we can write

$$I_j^{k+1} = \frac{E_j\left[\int_0^{\tau_k} r_{X(t)}\,dt\right]}{E_j[\tau_k]}.$$

Therefore it is possible to give a direct calculation of the index $I(j)$ of buffer j by

$$I(j) = \sup_{\tau \geq 0} \frac{E_j\left[\int_0^{\tau} r_{X(t)}\,dt\right]}{E_j[\tau]}, \tag{6}$$

where the supremum is taken over all stopping times τ of the Semi-Markov process $\{X(t)\}$ which are service completion times. The indices $I(j)$ are the so-called **Klimov indices**. From this observation we obtain the following properties of the indices.

Theorem 4. a) The indices computed by the largest remaining index algorithm fulfil

$$I^1 \geq I^2 \geq \ldots \geq I^N.$$

b) The Klimov indices fulfil

$$r_j \le I(j) \le \max_{1 \le i \le N} r_i, \quad j = 1, \dots, N.$$

c) If $\mu_1 = \ldots = \mu_N = 1$ and $i \mapsto r_i$ and $i \mapsto \sum_{j=1}^{k} p_{ij}$ are decreasing for all $k = 1, \dots, N$, then

$$I(1) \ge I(2) \ge \ldots \ge I(N).$$

Proof. a) and b) follow from Walrand [10] and from (6).

c) Let $K \in I\!R$, $0 < \beta < 1$ and define $V(j, K)$, $j = 1, \dots, N$ as the unique solution of the fixed point equation

$$V(j,K) = \max\{K, r_j + \beta \sum_{i=1}^{N} p_{ji} V(i,K)\}$$

Due to our assumptions about r and P we obtain that $V(j, K)$ is decreasing in j for all $K \in I\!R$. If we introduce now

$$I^\beta(j) = \min\{K \in I\!R \mid V(j,K) = K\},$$

it holds for $i \le j$ that $\{K \in I\!R \mid V(i,K) = K\} \subset \{K \in I\!R \mid V(j,K) = K\}$ and therefore $I^\beta(i) \ge I^\beta(j)$. Since $I(j) = \lim_{\beta \to 1}(1-\beta) I^\beta(j)$ the assertion is completely proved. □

4. Definition of the Priority Index Policy

The priority index policy is now defined as follows: assign the complete server to the non-empty buffer with highest index as long as there is fluid in this buffer. When the buffer is empty, assign to it only the capacity that is needed to hold the buffer at zero and assign the rest of the server to the buffer with second highest index and so on. Since there can be re-entrants from the newly processed buffer to buffers with higher priority, this procedures makes it necessary to re-assign the server capacity to all the buffers at each time point when a buffer empties. Before we define the index policy formally we assume that our network is strongly stable, i.e.

(A2) for all environment states $z \in Z$ there exists an $u = u(z) \in \mathcal{U}$ such that

$$\lambda^z < A^T u(z).$$

This condition simply means that it is possible to empty the system in finite time in each of the possible environment states.

Now fix an environment state $z \in Z$. An important implication of (A2) is that for a state y which fulfils $y_S = 0,\ y_j > 0,\ S \subset \{1, \ldots, N\},\ j \notin S$ it is possible to find an admissible allocation of the server such that

- there is no allocation to buffers $i \notin S + \{j\}$.
- the server is capable of keeping the buffers $i \in S$ at zero.
- the output rate at buffer j exceeds the input rate i.e. buffer j can be emptied.

This is because of the following observation: Define $T = S+\{j\}$ and $u_T = \lambda_T A_T^{-1}$. Hence $u_T \geq 0$ and $eu_T = e\lambda_T A_T^{-1} < 1$ by (A2). u_T is the server capacity that is needed to cope with the input at buffers $i \in T$. Now let $\varepsilon > 0$ and allocate an additional fraction of ε to buffer j. Due to re-entrants it is necessary to allocate some more capacity $v_S = \varepsilon a_j^S A_S^{-1} \geq 0$ to the buffers in S (see interpretation of the indices in the fluid setting). If we define $u^*_{S,j} = (u_T, 0) + (v_S, \varepsilon, 0)$ and choose $\varepsilon > 0$ such that $eu^*_{S,j} = 1$, we have found an admissible allocation which fulfils all requirements. In particular let us define the following important server allocations: $u^*_1 = (1, 0, \ldots, 0)$ and for $S_k = \{1, \ldots, k\}$, $u^*_k = u^*_{S_{k-1},k}$ as well as $u^*_0 = \lambda A^{-1}$.

Assume that the buffers have been rearranged such that the natural order coincides with the priority order i.e. $i_k = k,\ k = 1, \ldots, N$ and $S_k = \{1, \ldots, k\}$. Formally we will define the **priority index policy** as the stationary policy $\pi = (f, f, \ldots)$ with

$$f(y,z)(t) = \begin{cases} u^*_j(z) & \text{if} \quad j = \min\{i \mid y_i(t) > 0\} \\ u^*_0(z) & \text{if} \quad y_t = 0, \end{cases}$$

where $y_t = \phi^f(y,z)(t)$. To obtain the optimality we have to impose a further assumption which is very natural

(A3) the rewards $r_1, \ldots, r_N$ are non-negative.

Assumption (A3) guarantees that ideling of the server is not optimal as long as the buffers are nonempty, because processing of each buffer is profitable. Notice that the assumption $c_j \geq 0$ has not the same consequence. Obviously (A3) implies $c_j \geq 0$ whereas the converse is not valid.

Theorem 5 (Optimality of the priority index policy).

Under the assumptions $(A1) - (A3)$, the priority index policy is optimal for both the reward and the cost model of section 1.

The proof of the theorem is given in section 5. Hence, in the single-server network the optimal policy in the fluid model coincides with the optimal policy in the Semi-Markovian counterpart under the average reward criterion. It is remarkable that the priority index policy in the fluid model is independent of the discount factor β. This is mainly due to the fact that the impact of a decision occurs without delay. Some interesting **special cases** of the model are the following.

A) Zero-Routing

Let $P = 0$, which means that there is no routing and processed fluid leaves the system immediately. Obviously (A1) is fulfiled and (A2) reduces to

(A2') for all $z \in Z$ it must hold that $\sum_{j=1}^{N} \frac{\lambda_j^z}{\mu_j} < 1$.

Moreover, we obtain $r_j = \mu_j c_j$ and $a_j^S = 0$ for arbitrary $j \notin S$. Therefore the largest remaining index algorithm gives $I^{k+1} = \max_{j \notin S_k} r_j$ and the priority index policy is the well-known μc-rule.

B) Deteriorating Case

W.l.o.g. assume that $r_1 \geq r_2 \geq \ldots \geq r_N$ and let $p_{ji} = 0$ for $1 \leq i < j \leq N$. Hence P is an upper triangular matrix. This means that after processing a fluid it will only be routed to buffers with lower reward. In this case it is easy to see that $I^k = r_k$ and the index policy is the myopic or greedy policy.

C) Single-Server Re-Entrant Fluid Line

The following re-entrant fluid line is considered in Weiss [12]. Formally this problem reads

$$\begin{cases} \int_0^\infty e^{-\beta t} r u_t \, dt \to \max \\ \dot{y}_1(t) = \lambda - \mu_1 u_1(t) \\ \dot{y}_j(t) = \mu_{j-1} u_{j-1}(t) - \mu_j u_j(t), \quad j = 2, \ldots, N \\ y_t \geq 0 \\ u_t \in \mathcal{U} \end{cases}$$

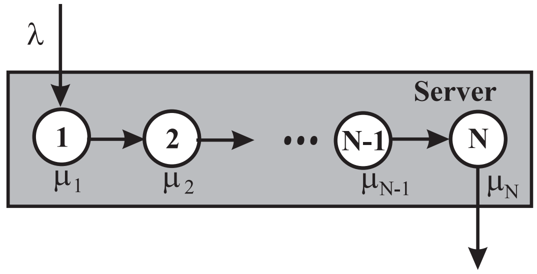

Figure 2. Single-server Re-Entrant Line

The routing matrix is given by

$$P = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & \dots & 0 \end{pmatrix}$$

Condition (A1) is not fulfiled in general in this model, but $(I - P_S)^{-1}$ obviously exists for all subsets $S \subset \{1, \dots, N\}$ which is sufficient for the derivation of all the preceeding lemmas and theorems. Let us define $m_j = \frac{1}{\mu_j}$. The stability condition (A2) reduces to $\lambda \sum_{j=1}^N m_j < 1$. Due to the special routing matrix the representation of the Klimov-indices in (6) gives

$$I(j) = \max_{1 \le t \le N-j+1} \frac{m_j r_j + \dots m_{j+t-1} r_{j+t-1}}{m_j + \dots + m_{j+t-1}}, \quad j = 1, \dots, N$$

which coincides with the indices given in Weiss [12] Proposition 3.2.

5. Proof of the optimality of the Index policy

In this section we show that the priority index policy is optimal for the defined control problems. We will prove this statement in the setting of the cost model, using a sample path argument.

In the sequel we will refer to the following deterministic control problem:

$$(C) \begin{cases} \int_0^\infty e^{-\beta t} c y_t \, dt \to \min \\ \dot{y}_t = \lambda_t - A^T u_t \\ y_t \ge 0 \\ u_t \in \mathcal{U} \end{cases}$$

where $\lambda_t \geq 0,\ t \geq 0$ is a given, deterministic function which is right continuous with left limits and has at most a countable number of discontinuities. The Hamiltonian of the control problem (C) is

$$H(y, u, p, t) = cy_t + p_t \lambda_t - u_t A p_t.$$

To identify an admissible control for (C) to be optimal, we can use the following lemma (cf. Seierstad/Sydsæter [7], where we use the weak regularity conditions).

Lemma 6 (Sufficient conditions for optimality).

The control u_t^* with the associate trajectory y_t^* is optimal for (C) if there exists a continuous and piecewise continuously differentiable vector function $p_t = (p_1(t), \ldots, p_N(t))$ as well as a piecewise continuous vector function $\eta_t = (\eta_1(t), \ldots, \eta_N(t))$ such that for all $t \geq 0$

(i) u^* maximizes $u_t \mapsto u_t A p_t$ for $u_t \in \mathcal{U}(\dagger_\sqcup, \ddagger)$.

(ii) $\dot{p}_t - \beta p_t = -c + \eta_t$, where the $p_j(t)$ are differentiable, $j = 1, \ldots, N$.

(iii) $\eta_t \geq 0$.

(iv) $\eta_t y_t^* = 0$.

(v) $\liminf_{t\to\infty} e^{-\beta t} p_t (y_t^* - y_t) \geq 0$ for all admissible trajectories y_t.

The next lemma will be important in the proof of optimality. Notice, that the statement of Lemma 7 was also used in Chen/Yao [4] as a further condition for global optimality of the index policy. The authors there have not seen that this condition is always fulfiled. Hence our lemma also fills a gap in Chen/Yao [4].

Lemma 7. For $k = 1, \ldots, N-1$ it holds that

$$A_{S_k}^{-1}(I^{k+1} e_{S_k} - r_{S_k}) \leq 0.$$

Proof. W.l.o.g. we assume that $i_k = k$ for $k = 1, \ldots, N$. By v^j we denote the unit vector with 1 in component j. The dimension should be clear from the context. We have to show that

$$I^{k+1} \leq \frac{v^j A_{S_k}^{-1} r_{S_k}}{v^j A_{S_k}^{-1} e_{S_k}} \quad \text{for all } 1 \leq j \leq k, \ \ k = 1, \ldots, N.$$

We will do this by induction over k.

For $k = 1$ we have $I^1 = r_1$ and the inequality obviously holds. Assume the statement is valid for $I^1, \ldots, I^k$. We define $w_j = a_j^{S_{k-1}} A_{S_{k-1}}^{-1}$, $k \le j \le N$ and

$$A_{S_k} = \begin{pmatrix} A_{S_{k-1}} & a_S \\ -a_k^{S_{k-1}} & a_{kk} \end{pmatrix}$$

Hence we have

$$I^k = \frac{r_k + w_k r_{S_{k-1}}}{1 + w_k e_{S_{k-1}}}.$$

Define the k-dimensional vector $z^k = v^k A_{S_k}^{-1}$, hence

$$z^k = \begin{pmatrix} w_k \alpha \\ \alpha \end{pmatrix} \quad \text{where} \quad \alpha = (a_S w_k + a_{kk})^{-1} \neq 0.$$

This can be easily checked by computing $A_{S_k}^T z^k$. This gives us

$$\frac{v^k A_{S_k}^{-1} r_{S_k}}{v^k A_{S_k}^{-1} e_{S_k}} = \frac{z^k r_{S_k}}{z^k e_{S_k}} = \frac{r_k \alpha + \alpha w_k r_{S_{k-1}}}{\alpha + \alpha w_k e_{S_{k-1}}} = I^k$$

and by Theorem 4 a) $I^{k+1} \le I^k$. What is left to prove is that

$$I^{k+1} \le \frac{v^j A_{S_k}^{-1} r_{S_k}}{v^j A_{S_k}^{-1} e_{S_k}} \quad \text{for all } 1 \le j \le k-1.$$

Let $z^j = v^j A_{S_k}^{-1}$, hence it is easy to check that

$$z^j = \begin{pmatrix} w_k \alpha + v^j A_{S_{k-1}}^{-1} \\ \alpha \end{pmatrix} \quad \text{where} \quad \alpha = \frac{-a_S v^j A_{S_{k-1}}^{-1}}{a_{kk} + a_S w_k}.$$

Therefore

$$\frac{v^j A_{S_k}^{-1} r_{S_k}}{v^j A_{S_k}^{-1} e_{S_k}} = \frac{z^j r_{S_k}}{z^j e_{S_k}} = \frac{\alpha \left(r_k + w_k r_{S_{k-1}} \right) + v^j A_{S_{k-1}}^{-1} r_{S_{k-1}}}{\alpha \left(1 + w_k e_{S_{k-1}} \right) + v^j A_{S_{k-1}}^{-1} e_{S_{k-1}}} \ge$$

$$\frac{r_k + w_k r_{S_{k-1}}}{1 + w_k e_{S_{k-1}}} = I^k \quad \Leftrightarrow \quad \frac{v^j A_{S_{k-1}}^{-1} r_{S_{k-1}}}{v^j A_{S_{k-1}}^{-1} e_{S_{k-1}}} \ge \frac{r_k + w_k r_{S_{k-1}}}{1 + w_k e_{S_{k-1}}} = I^k$$

which is true by the induction hypothesis. Since $I^k \ge I^{k+1}$ (see Theorem 4 a)) we obtain the desired result. □

Now we are able to prove the main theorem.

Proof of Theorem 5:. We will prove Theorem 5 for the cost problem by using a sample path argument: Let π be the index policy as defined before. We have already shown that $\pi \in F^\infty$. Suppose $\sigma \in F^\infty$ is an arbitrary policy. We have to show that $C_\pi(y,z) \le C_\sigma(y,z)$ for $(y,z) \in I\!R_+^N \times Z$. Now let $\omega \in \Omega$ be fixed. Denote by $\lambda_t = \lambda^{Z(t,\omega)}$ the path of the input process, by $\pi(t,\omega)$, $\sigma(t,\omega)$ the paths of the controls and by $Y^\pi(t,\omega)$ $(Y^\sigma(t,\omega))$ the path of the buffer contents under control π (σ). Due to the definition of the deterministic control problem (C), $\pi(t,\omega)$ and $\sigma(t,\omega)$ are admissible controls for (C), with respective costs $\int_0^\infty cY^\pi(t,\omega)\,dt$ and $\int_0^\infty cY^\sigma(t,\omega)\,dt$.

By showing that there exist functions p_t and η_t which fulfil together with $\pi(t,\omega)$ and $Y^\pi(t,\omega)$ the sufficient conditions of Lemma 3 we prove that

$$\int_0^\infty cY^\pi(t,\omega)\,dt \;\le\; \int_0^\infty cY^\sigma(t,\omega)\,dt \tag{7}$$

for all $\omega \in \Omega$. Integrating w.r.t. ω on both sides we finally obtain the desired inequality

$$C_\pi(y,z) \le C_\sigma(y,z).$$

In what follows we will suppress ω.

Before we start to prove (7), let us introduce several notations: let

$$R = (r,\ldots,r) \text{ and } C = (c,\ldots,c) \text{ be } N \times N\text{-matrices.}$$

By $0 = t_0 \le t_1 \le t_2 \le \ldots t_N < \infty$ denote the successive emptiness times of the buffers under the priority index policy π. Notice that such a relation of the emptiness times can be guaranteed under the index policy, regardless of the choosen $\omega \in \Omega$. (Of course the t_j depend on ω). Some of the buffers may initially be empty, in which case $0 = t_1 = \ldots = t_i$ holds. Denote

$$T = \begin{pmatrix} -e^{-\beta t_1} & 0 & \ldots & 0 \\ e^{-\beta t_1} - e^{-\beta t_2} & -e^{-\beta t_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ e^{-\beta t_{N-1}} - e^{-\beta t_N} & e^{-\beta t_{N-1}} - e^{-\beta t_N} & \ldots & -e^{-\beta t_N} \end{pmatrix}$$

For $j = 1,\ldots,N$ we define now the adjoint functions p_j in the following way

$$p_j(t) = 0 \quad \text{when } t \ge t_N,$$
$$p_j(t) = \frac{1}{\beta}\left(b_{jk} + l_{jk}e^{\beta t}\right) \quad \text{when } t_{k-1} \le t < t_k,\ k = 1,\ldots,N,$$

where $b_{jk} = c_j$ if $k \leq j$ and the other constants will be determined as the proof continues. Hence, condition (ii) is obviously satisfied with $\eta_j(t) = 0$ if $t \leq t_j$ which also implies that (iv) is true. Moreover, the adjoint functions are piecewise continuously differentiable and the Lagrange multipliers η_j are constant on the intervals $[t_k, t_{k+1})$. To ease notation, let us define

$$B = (b_{jk}) = \begin{pmatrix} c_1 & b_{12} & b_{13} & \cdots & b_{1N} \\ c_2 & c_2 & b_{23} & \cdots & b_{2N} \\ \vdots & \vdots & \vdots & & \vdots \\ c_N & c_N & c_N & \cdots & c_N \end{pmatrix}$$

$$L = (l_{jk}) = \begin{pmatrix} l_{11} & l_{12} & l_{13} & \cdots & l_{1N} \\ l_{22} & l_{22} & l_{23} & \cdots & l_{2N} \\ \vdots & \vdots & \vdots & & \vdots \\ l_{NN} & l_{NN} & l_{NN} & \cdots & l_{NN} \end{pmatrix}$$

What is left to prove, are conditions (i) and (iii) as well as the continuity of p. We will now determine the constants l_{jk} in such a way that p is continuous. The continuity condition for $1 \leq j \leq k < N$ reads

$$\begin{aligned} b_{jN} + l_{jN} e^{\beta t_N} &= 0 \\ b_{jk} + l_{jk} e^{\beta t_k} &= b_{jk+1} + l_{jk+1} e^{\beta t_k}. \end{aligned}$$

From this set of equations, we obtain the following recursion for each $j = 1, \dots, N$:

$$\begin{aligned} l_{jN} &= -b_{jN} e^{-\beta t_N} \\ l_{jk} &= l_{jk+1} + (b_{jk} - b_{jk+1})(-e^{-\beta t_k}) \quad k = N-1, \dots, j. \end{aligned}$$

Hence, we get inductively that

$$l_{jk} = \sum_{i=k}^{N-1} (b_{ji} - b_{ji+1})(-e^{-\beta t_i}) - b_{jN} e^{-\beta t_N}$$

for $1 \leq j \leq k \leq N$. Using matrix notation this is simply $L = BT$. Hence $L = BT$ guarantees that p is continuous and

$$p_j(t) = \frac{1}{\beta} \left[(B)_{jk} + (BT)_{jk} e^{\beta t} \right], \quad \text{when } t_{k-1} \leq t < t_k,$$

Now we are going to show that (i) can be satisfied by choosing the remaining constants b_{jk} appropriately. For that purpose define $R_t = (R_1(t), \ldots, R_N(t)) = Ap_t$. To fulfil condition (i) we have to show that

$$0 \leq R_1(t) = \ldots = R_j(t) > R_{j+1}(t), \ldots, R_N(t) \quad \text{when } t_{j-1} \leq t < t_j$$

for $j = 1, \ldots, N$ and

$$R_1(t) = \ldots = R_N(t) = 0 \quad \text{when } t \geq t_N.$$

Since $p_t = 0$ for $t \geq t_N$ the last equation trivially holds. Due to the definition of the adjoint functions we can write

$$R_j(t) = \frac{1}{\beta}\left[(AB)_{jk} + (ABT)_{jk} e^{\beta t}\right], \quad \text{when } t_{k-1} \leq t < t_k,$$

Hence we get

$$R_j(t) = \frac{1}{\beta}\left(AB[I + Te^{\beta t}]\right)_{jk}, \quad \text{when } t_{k-1} \leq t < t_k.$$

For abbreviation denote $X = (x_{jk}) = AB$. Since for $t \in [t_{k-1}, t_k)$ it holds that the k-th column of $(I + Te^{\beta t})$ is non-negative, it is not difficult to see that for (i) it suffices to show that

$$\begin{aligned}
&0 \leq x_{11} \geq x_{21}, \ldots, x_{N1} \\
&0 \leq x_{12} = x_{22} \geq x_{32}, \ldots, x_{N2} \\
&\quad \vdots \\
&0 \leq x_{1N} = \ldots = x_{NN}.
\end{aligned}$$

If we define $\hat{B} = B - C$ we have $X = A(\hat{B} + C) = A\hat{B} + R$. Since $\hat{b}_{j1} = 0,\ j = 1, \ldots, N$ it follows that $x_{j1} = r_j,\ j = 1, \ldots, N$. Hence the first assertion follows from the definition of the first index and the fact that $r_1 \geq 0$. Let us now have a look at column $k+1,\ k \in \{1, \ldots, N-1\}$ of matrix X. With $\hat{b}_{k+1} = (\hat{b}_{1,k+1}, \ldots, \hat{b}_{k,k+1})$ and $S_k = \{1, \ldots, k\}$ define

$$\hat{b}_{k+1} = A_{S_k}^{-1}\left(I^{k+1} e_{S_k} - r_{S_k}\right).$$

Hence we obtain that

$$(x_{1,k+1}, \ldots, x_{N,k+1}) = \begin{pmatrix} I^{k+1} e_{S_k} \\ a_{k+1}^{S_k} A_{S_k}^{-1} r_{S_k} - I^{k+1} a_{k+1}^{S_k} A_{S_k}^{-1} e_{S_k} + r_{k+1} \\ \vdots \\ a_N^{S_k} A_{S_k}^{-1} r_{S_k} - I^{k+1} a_N^{S_k} A_{S_k}^{-1} e_{S_k} + r_N \end{pmatrix}$$

Because of the definition of I^{k+1}, it holds that

$$x_{k+1,k+1} - I^{k+1} = \left(r_{k+1} + a_{k+1}^{S_k} A_{S_k}^{-1} r_{S_k}\right) - I^{k+1}\left(1 + a_{k+1}^{S_k} A_{S_k}^{-1} e_{S_k}\right) = 0$$

and since $a_i^{S_k} \geq 0$, $A_{S_k}^{-1} \geq 0$, using the maximality of I^{k+1} we have for $i = k+2, \dots, N$:

$$x_{i,k+1} - I^{k+1} = \left(r_i + a_i^{S_k} A_{S_k}^{-1} r_{S_k}\right) - I^{k+1}\left(1 + a_i^{S_k} A_{S_k}^{-1} e_{S_k}\right) \leq 0$$

and using additionally (A3) we get $I^{k+1} \geq 0$. Hence we have shown that $0 \leq x_{1,k+1} = \ldots = x_{k+1,k+1} \geq x_{k+2,k+1}, \dots, x_{N,k+1}$, $k \in \{1, \dots, N-1\}$ which implies (i) of the sufficient conditions.

Due to Lemma 7 it holds that $\hat{b}_{k+1} \leq 0$, $k = 1, \dots, N-1$ or equivalently $B \leq C$. Bearing condition (ii) in mind, this implies that the Lagrange multipliers η_t are non-negative, hence (iii) holds. Since $t_N < \infty$ we obtain that $p_t = 0$ for all $t > t_N$ which implies (v). This completes the proof. □

6. Conclusion

We have shown that the optimal policy for a stochastic single-server fluid network is simple priority index policy. The indices coincide with the optimal indices in a Semi-Markovian Klimov problem. In particular are these indices independent of the environment process (Z_t) and of the discount factor β. From the sample path proof of Theorem 4 it can be seen that the index policy is optimal in a very strong sense: it stochastically minimizes the cost function over all time intervals $[0,t),, t \in I\!R_+ \cup \{\infty\}$. Thus, this policy can also be interpreted as a myopic policy. From that point of view the index policy is quite natural. Moreover, the sample path proof shows that, even one were knowing in advance the complete sample path of the environment process, the optimal policy would remain the same.

References

[1] F. Avram, Optimal Control of Fluid Limits of Queueing Networks and Stochasticity Corrections, in *Mathematics of Stochastic Manufacturing Systems*, eds. G.G. Yin and Q. Zhang, Lect. Appl. Mathem. **33** (1997) 1-36.

[2] F. Avram, D. Bertsimas, M. Ricard, Fluid Models of Sequencing Problems in open Queueing Networks: an Optimal Control Approach, in *Stochastic Networks*, eds. F.P. Kelly and R.J. Williams (1995) 199-234.

[3] H. Chen and A. Mandelbaum, Hierarchical modeling of stochastic networks part 1: fluid models, in *Stochastic modeling and analysis of manufacturing systems*, ed. D.D. Yao, Springer-Verlag, New York (1994) 47-105.

[4] H. Chen and D.D. Yao, Dynamic Scheduling of a Multiclass Fluid Network, Operations Research **41** (1993) 1104-1115.

[5] G.P. Klimov, Time-sharing service systems I, Th. Prob. Appl. **19** (1974) 532-551.

[6] S. Meyn, Stability and optimization of queueing networks and their fluid models, in *Mathematics of Stochastic Manufacturing Systems*, eds. G.G. Yin and Q. Zhang, Lect. Appl. Mathem. **33** (1997) 175-199.

[7] A. Seierstad and K. Sydsæter, *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam (1987).

[8] S.P. Sethi and Q. Zhang *Hierarchical decision making in stochastic manufacturing systems*, Birkhäuser, Boston (1994).

[9] D.W. Tcha and S.R. Pliska, Optimal control of single server queueing networks and multi-class $M/G/1$ queues with feedback, Oper. Res. **25** (1977) 248-258.

[10] J. Walrand, *An Introduction to Queueing Networks*, Prentice-Hall, Englewood Cliffs, New Jersey (1988).

[11] J. Weishaupt, Optimal myopic policies and index policies for stochastic scheduling problems, ZOR-Mathem. Meth. Oper. Res. **40** (1994) 75-89.

[12] G. Weiss, On Optimal Draining of Re-entrant Fluid Lines, in *Stochastic Networks*, eds. F.P. Kelly and R.J. Williams (1995) 91-103.