**Author**: Ertiningsih, D.

**Title**: Structural Properties of Single Server Queueing Systems: Efficient Methods via Lumping and Dynamic Programming

**Issue Date**: 2020-02-05

# Chapter 1

# Introduction

In this thesis we will consider random, discrete-valued queueing systems, and in particular structural properties of performance measures of interest that are based on their long run average behaviour. In practice, the state space of such a discrete-valued random process is finite, but may be large. However, the analysis, both mathematically and numerically, of such a system is generally hard. First of all, as the state space gets large, computational complexity hampers one's ability to draw valid inference for the problems of interest. Secondly, potential structural properties are affected by the existence of state space boundaries.

It may therefore be beneficial to consider the finite state space model as an approximation or a perturbation of the infinite state space model. Systematic properties of the infinite state space model can then be exploited to obtain insight in the properties of the finite state space model, and to develop more efficient algorithms for its numerical study. Thus, the object of this thesis are countable state queueing systems. For simplicity we will restrict to the single server case.

The main focus of this thesis is to study structural properties of the queueing system at hand. Structural properties such as a Quasi-Birth-Death structure or a Quasi-Skipfree structure provide insight in the form of the stationary distribution.

Structural properties play a role as well, when control is exercised in order to optimise a performance measure of interest. When the controlled system is a discrete time Markov decision process (MDP), value iteration (VI) is an important algorithm for the practical analysis of such a process. It is not only used for numerical analysis, but also for the derivation of structural properties of the value function, such as increasingness, convexity, supermodularity, in order to deduce structural properties of optimal policies. In the next sections, we will give an overview of the results in this thesis. The embedding in the existing literature will be elaborated upon in the appropriate chapters.

## 1.1 Stationary distribution of single server queues

In the first part of this thesis consisting of Chapters 2 and 3, we are interested in analysing the stationary distribution of a single server queueing system that can be modelled as a Markov process. This stationary distribution can be used to derive, for instance, the average number of customers in a queueing system. The single server queue is visualised in Figure 1.1.

We will first illustrate our approach in these two chapters by an analysis of the GI/M/1 queue. Then, we discuss the main results from each chapter.

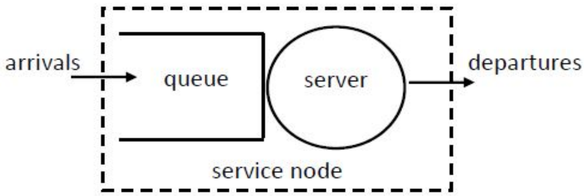


Figure 1.1: Single server queue

The stationary distribution is well-known to have a product form in the number of customers present in the system, when the arrival process is a Poisson process. If the arrival process is a Poisson process, the stationary distribution is well-known. In the case of a general arrival process with i.i.d. interarrival times, the stationary distribution is a modified geometric distribution (cf. Asmussen [5]) with a parameter that is defined by an *implicit* function relation. Since Cox distributions are dense in the space of non-negative distributions (cf. [10]), we can approximate any interarrival distribution arbitrarily closely by a Cox distribution.

Thus, assuming the interarrival times to have a Cox distribution, one can use Markov process techniques to show that the parameter of the modified geometric distribution is the solution to an *explicit* eigenvalue problem. For simplicity of the exposition, in this introduction we will assume the interarrival times to have an Erlang distribution with $k$ exponentially distributed phases, each with parameter $\lambda$.

This leads to the $E_k/M/1$ queueing system. The service times are assumed to be exponentially distributed with parameter $\mu$, and customer interarrival times and service times are assumed independent. This system can be modelled as a Markov process, if we include the number of phases left till the next arrival, in our state description. This yields a two-dimensional state space $\boldsymbol{S} = \{(m,i) \,|\, m \in \mathbf{Z},\, i \in \{0, \cdots, k-1\}\}$, where $m$ is called the 'level' of the state and $i$ denotes the 'phase' within the level.

As a result, since the rates are independent of the level except for the one at the left end, the above $E_k/M/1$ queueing system can be modelled as a (time) homogeneous Quasi-Birth-Death (QBD) process $X = \{X(t)\}_{t\geq 0}$ on the state space $\boldsymbol{S} = \{L_0, L_1, \cdots\}$, where $L_m = \{(m,0),(m,1),...,(m,k-1)\}$, for all $m \geq 0$. The process is a QBD process, because the jump rates from level $m$ do not cross more than one level in both directions, i.e., positive jump rates only occur to levels $(m-1)^+ = \max(0, m-1), m$, and $(m+1)$. The graphical representation of this process is given in the Figure 1.2. Thus, the transition rate matrix $Q$ takes the form

$$Q = \begin{bmatrix} W_0 & U & 0 & 0 & 0 & \cdots \\ D & W & U & 0 & 0 & \cdots \\ 0 & D & W & U & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where $W_0, W, U$, and $D$ are all $k \times k$ matrices denoting the transition rates within a level $(W_0, W)$, to the higher next level $(U)$, and to the next lower level $(D)$. The transition rate matrix $Q$ is a conservative, irreducible, and stable $q$-matrix. Further, we assume that the associated QBD process is ergodic, which in this case means that $\lambda/k\mu < 1$.

The stationary distribution $\pi$ of homogeneous QBD processes can be determined via the matrix analytic method (see e.g. [29]) as follows:

$$\pi_m = \pi_{m+1}\, R_m, \;\; R_m = D\, T_m, \;\; m \geq 0, \tag{1.1.1}$$

where $\pi_m$ is the stationary distribution restricted to the states of level $m$, and $T_m$ represents the expected amount of time spent in level $m$ without passing through levels $n > m$, given that the system starts in level $m$.

As Figure 1.2 shows for the $E_k/M/1$ queueing system, each level can be entered from the left levels only in one state, called an *entrance state*. Note that, by the entrance state property, $T_m$ is independent of the rates within lower levels, $m \geq 1$, as long as the associated Markov process is irreducible. The existence of an entrance state allows us to compute the stationary distribution via the so-called *successive lumping* method, which has been introduced by Katehakis and Smit [23]. We will explain this method briefly next.

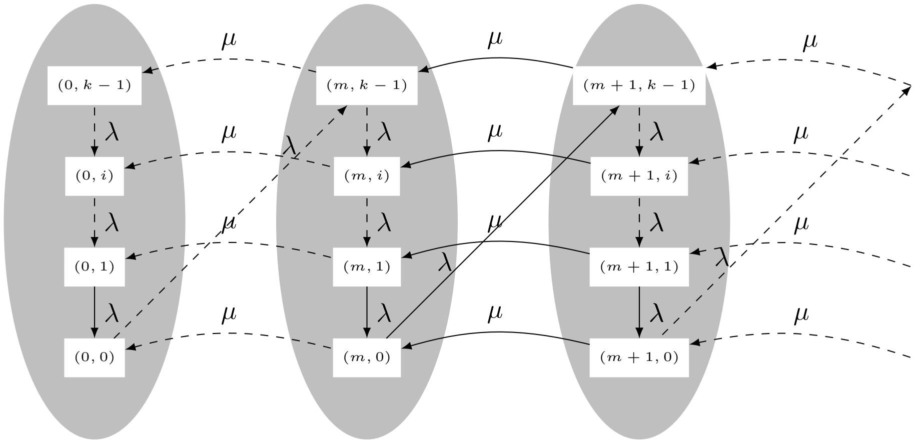


Figure 1.2: $E_k/M/1$ queueing system

Successive lumping is a general procedure to calculate the stationary distribution of an irreducible and ergodic Markov process $X$ on the state space $\boldsymbol{S}$. To use this procedure, $\boldsymbol{S}$ is partitioned into sets $L_0, \cdots, L_N$, where $N \leq \infty$. These sets are called *levels*, and the partition is called a *level partition* (cf. Chapters 2 and 3). The successive lumping procedure requires the presence of so-called *entrance states*. A set $A \subsetneqq X$ has an *entrance state*, $x \in A$ say, if the one-step transitions from states $y \in \boldsymbol{S} \backslash A$ to $A$ always end in $x$. The Markov process $X$ is called *successively lumpable with respect to the state space partition* $\boldsymbol{S} = \cup_j L_j$ if either $\underset{\sim}{L}_n := \cup_{j \leq n} L_j$ or $\widetilde{L}_n := \cup_{j \geq n} L_j$ has an entrance state for all $n$.

The procedure to compute the stationary distribution via successive lumping is as follows. Assume that $\underset{\sim}{L}_n$ has an entrance state for all $n$. First, we compute the relative stationary distribution in the first level. Then, the first level is lumped as a single state and combined together with the second level. Next, we compute the relative stationary distribution in that new combined level by using the first result. This procedure is continued for the remaining levels. Finally, we combine all results and compute the stationary distribution of the system. See [23] and [24] for more details on successive lumping, a formal definition of entrance states and DES (down entrance state) processes. More specifically, we refer to Definition 1 of [23] and Lemmas 1 and 2 of [24]. For the usage of successive lumping to find rate matrices within the matrix analytic framework, we refer to [24].

In case $L_0, \cdots, L_N$ are all finite, and $\underset{\sim}{L}_n$ has an entrance state for all $n$, the successive lumping procedure is efficient, even if $N = \infty$. If, on the contrary, $\widetilde{L}_n$ has an entrance state for all $n$, $L_0, \cdots, L_N$ are all finite, and $N = \infty$, then this is not necessarily the case. However, if the level sets correspond to the level sets of a QBD, such as the $E_k/M/1$ queueing system, the successive lumping method has a simpler implementation yielding Eqn. (1.1.1) with $T_m$ and $R_m$ explicitly computable and independent of $m$, for $m \geq 1$.

In particular, in the case of the $E_k/M/1$ queueing system, the matrix $D$ in Eqn. (1.1.1) is a diagonal matrix, and hence invertible. It then holds, that

$$\pi_{m+1} = -\pi_m \check{Q} D^{-1}, \quad m \geq 1, \tag{1.1.2}$$

where the matrix $\check{Q}$ is given by

$$\check{Q} = \begin{bmatrix} -(\lambda+\mu) & 0 & \cdots & 0 & \mu \\ \lambda & -(\lambda+\mu) & \cdots & 0 & \mu \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -(\lambda+\mu) & \mu \\ 0 & 0 & \cdots & \lambda & -\lambda \end{bmatrix}.$$

The analysis in Chapter 2 yields validity of the following formula

$$\pi_{(mj)} = \pi_{(00)} \gamma^m \beta_1^j, \quad m \geq 1,$$

with $\gamma = 1/\beta_1^k$ the minimum eigenvalue in absolute value of $-\check{Q}D^{-1}$. In other words, the stationary distribution is a level product form distribution.

Next, to give an explicit formula for the stationary distribution of the $E_k/M/1$ queue we need to compute the stationary probabilities $\pi_0 = (\pi_{(00)}, \pi_{(01)}, ..., \pi_{(0,k-1)})$ of the boundary states, i.e. of level $L_0$.

This yields the following result

$$\pi_{(mj)} = \begin{cases} \dfrac{(1-\gamma)\gamma^m \beta_1^j}{k}, & m > 0 \\ \dfrac{1-\gamma\beta_1^j}{k}, & m = 0, \end{cases}$$

where $j = 0, 1, \ldots, k-1$.

By taking the sum over the phases per level, we get the well-known formula for the stationary probability $\bar{\pi}_m = \sum_{j=0}^{k-1} \pi_{mj}$ of $m$ customers in the $E_k/M/1$ queue (cf. Asmussen [5] Theorem 5.1)

$$\bar{\pi}_m = \begin{cases} (1-\gamma)\rho\gamma^{m-1}, & m > 0 \\ 1-\rho, & m = 0, \end{cases}$$

where $\rho = \lambda/k\mu$.

We would like to mention that the successive lumping method is applicable to more complex networks. For instance, customers may renege at an exponential rate. This yields a non-homogeneous QBD. Another example is, that customers arrive in random batches after an exponential lead time. In this case, the process may jump from a given state to non-neighbouring levels on the right. Thus, we obtain a so-called Quasi-Skipfree Process (QSF)(cf. Latouche and Ramaswami [29]).

The above model has already been analysed by Latouche and Ramaswami [32], in discrete time, not from the entrance state perspective, but by requiring that the upward matrix $U$ has rank 1. The analysis in that paper is based on the rate matrix $R'_m = UT'_m$, where $T'_m$ is the sojourn time spent in level $m$ without passing through the lower levels $0, .., m-1$, $m \geq 1$, in contrast with our approach. Then, the relation $\pi_{m+1} = \pi_m R'_m$, $m \geq 0$, holds.

If $U$ has rank 1, then trivially $R'_m$ has rank 1 and $\gamma$ is the unique nonzero eigenvalue of $R'_m$. Notice, that by homogeneity $R'_m$ is independent of $m$. However, the rate matrix cannot be explicitly computed and the eigenvalue $\gamma$ can only be determined as a solution to a fix point equation. This is in fact a general result for homogeneous QBDs with a rank 1 upward matrix $U$, proved in [32].

The approach sketched above for the $E_k/M/1$ queueing system, mainly concerns Chapter 2. In Chapter 2, we study QSF processes that are skipfree to the right, with a rank 1 upward matrix (in the set-up of this introduction!). As an extension of the result by Ramaswami and Latouche [32], we prove that the stationary distribution has a level product form in $\gamma < 1$, say. If the QSF process is a QBD process, and the upward matrix has one non-zero column (entrance state structure) or one non-zero row (exit state structure), then the matrix $R_m$ in Eqn. 1.1.1 can be explicitly computed, and $1/\gamma$ is shown to be the maximum eigenvalue (in absolute value) of this matrix. In the particular case of an invertible down matrix $D$, one does not need to solve an eigenvalue problem. Instead, we get the explicit relation 1.1.2, expressing $\pi_{m+1}$ in terms of $\pi_m$. This allows to solve $\pi$ directly.

In Chapter 3 our main concern is the question, whether upward (or downward) rank 1 matrices are more general than upward matrices with either one non-zero column or one non-zero row. We show that this is not the case. By introducing an extra state per level, we construct a perturbed system with the added states being upward level entrance states. We show that the stationary distribution of the perturbed system converges to the stationary distribution of the original model, as the perturbation parameter tends to 0. In the limit to 0 of the perturbation parameter, these added states become instantaneous entrance states.

**Open problems**

Many questions have to be resolved yet. Below we list a few imminent ones:

- Is it possible to extend the successive lumping procedure to the case of more than one entrance state? Could one achieve this by the use of instantaneous states?

- Suppose that the process is killed at an exponential rate. Can one use successive lumping to compute the expected time spent in each state before the process is killed? This could be of interest for computing the discounted value function. A first analysis in this direction, without using successive lumping, can be found in [35], where the priority queueing system with two priority classes is studied.

## 1.2 Structural properties of the value function and optimal control of single server queues

In the second part of this thesis, we are interested in structural properties of the value function, such as monotonicity, convexity, and supermodularity. These properties allow to infer that there exist threshold or switching curve optimal policies, depending on whether the state space has dimension one or two respectively.

In classical queueing models, it is assumed that customers are willing to wait until getting served. However, in real life, customers may lose patience due to long service times or long queues, or more compelling reasons, and so they may leave without having completed service. For instance, Armony, Plambeck, and Seshadri [4] have studied the multi-server queueing model with impatient customers. They apply sample path arguments to derive convexity properties for an unobservable $M/M/s$ queue with impatient customers. Therefore, in modern controlled queueing models, it is important to allow for customer impatience (or abandonment) and retrial.

Due to customer abandonment, the transition rates are unbounded, and so standard mathematical tools for deriving the desired structural properties cannot be used. Indeed, the analysis is hindered by the fact that the system is not uniformisable, so that the usual discrete time Markov decision process algorithms cannot be applied. A further complicating factor is that structural properties are generally lost by a standard state space truncation (cf. [16]). In [11], Bhulai, Brooms, and Spieksma have developed the so-called Smoothed Rate Truncation (SRT) method. Application of this truncation method generally leaves the structural properties of the value function of the infinite system intact. So far, this method has been shown to apply only to Markov processes satisfying strong stability conditions. Usually processes with customer abandonment satisfy these strong conditions. However, in general, exponential distributions do not model service time distributions accurately. This leads to a non-Markovian controlled system, and so the tools that already have been mentioned above, cannot be applied. So far, no general method for addressing these issues seems to have been developed.

In Chapter 4, we aim to develop a coupling method that allows to derive structural properties of the value function of non-uniformisable systems without restricting to exponential transition times. The main model illustrating this technique is the non-controlled $GI/G/1+M+G/\text{FIFO}$ queueing system. Thus, the external arrival process to a single server queue is a renewal process, the service time distribution is allowed to have a general distribution, and customers may abandon the queue or the server after an exponentially distributed amount of time. After abandonment, a customer decides (according to a given probability distribution) to either leave the system, or to retry after an amount of time with a general distribution. Our coupling method does not yet allow a general abandonment distribution. Arrival process, service times, abandonment times and retrial times are all assumed independent.

The stochastic process associated with this queueing system records at any time the number of customers at the single server system as well as the number of customers that have abandoned and are waiting for retrial. Thus, this stochastic process has a two-dimensional state space. We use our coupling method to show that the total cost incurred till time $t$ is non-decreasing and convex in both state-variables, as well as supermodular provided the holding cost per unit time has these properties, as a function of the states.

To illustrate the basic idea of our coupling construction, we consider the simpler $GI/M/1+M$ queueing system in this introduction. The representation of this system is given in Figure 1.3. We will also specify our coupling slightly differently from Chapter 4 for an easier understanding. The state $X_t$ of the system at time $t$ denotes the number of customers present at time $t$. $c(x)$ is the incurred cost rate when $X_t = x$, and we assume that $x \mapsto c(x)$ is a convex, non-decreasing function on $\{0, 1, \cdots\}$. Denote by $V_t(x) = \mathbb{E}\left[\int_0^t c(x_u)du \mid x_0 = x\right]$ the total cost incurred till time $t$, given that the initial state equals $x$. The objective is to show that $x \mapsto V_t(x)$ is a convex, non-decreasing function on $\{0, 1, \cdots\}$.

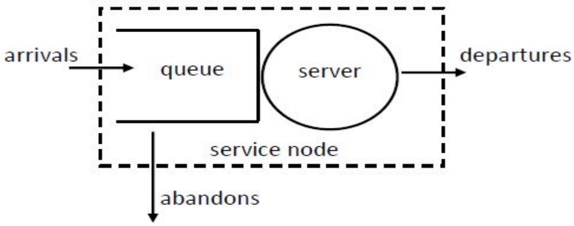


Figure 1.3: Scheme of the $GI/M/1+M$ queue

To this end, fix $x$. We construct four processes $X^i = (X^i_t)_{t\geq 0}$, for $i = 1, 2, 3, 4$, satisfying

$$X^1_0 = x,\ X^2_0 = X^3_0 = x+1,\ X^4_0 = x+2$$

and

$$0 \leq X^2_t - X^1_t \leq X^4_t - X^3_t,\ t \geq 0, \tag{1.2.1}$$

such that for all $B \subset \{0, 1, \cdots\},\ t \geq 0$

$$\begin{aligned} \mathbb{P}(X_t \in B \mid x_0 = x) &= \mathbb{P}\left(X^1_t \in B\right), \\ \mathbb{P}(X_t \in B \mid x_0 = x+1) &= \mathbb{P}\left(X^2_t \in B\right) = \mathbb{P}\left(X^3_t \in B\right) \\ \mathbb{P}(X_t \in B \mid x_0 = x+2) &= \mathbb{P}\left(X^4_t \in B\right). \end{aligned}$$

The above conditions imply that the set of random variables $\{X^1_t, X^2_t, X^3_t, X^4_t\}$ is stochastically increasing and convex in the sense of [38, Definition 3.1] for $t \geq 0$. By [38, Proposition 3.2], and by subsequently out-integrating over time, non-decreasingness and convexity of the function $x \mapsto V_t(x)$, $t \geq 0$, follows.

To achieve this, we draw independent sequences of (independent) interarrival times, service times, and abandonment times. Events are epochs of a new arrival generated by the arrival process, a service completion, or an abandonment. At each event, the customers at the server in their respective processes get the next service time realisation (this is feasible, because the service times have an exponential distribution).

If the event is due to an arrival, we proceed as follows. A new customer is assigned to each process and in each of these, he is assigned an abandonment time from the generated sequence, say $\{I_n\}_n$. This assignment has the following property: at any time $t \geq 0$, for any $n \in \{1, 2, \ldots\}$, precisely one customer in

(1) each of the four processes, or only in

(2) processes $X^2$ and $X^4$ or

(3) processes $X^3$ and $X^4$, or

(4) process $X^4$

has been assigned the $n$-th abandonment time $I_n$, or no customer in any of the four processes has been assigned $I_n$. For independence reasons, $I_n$ can be assigned at most once per process to a customer.

As an example, at time 0, $I_1, ..., I_x$ have been assigned to customers in all four processes, $I_{x+1}$ to a customer in processes $X^2$ and $X^4$, and $I_{x+2}$ to a customer in processes $X^3$ and $X^4$. To achieve that the required properties on the assigned abandoment times hold at any time $t$, requires some delicate handling. It easily follows that Eqn. 1.2.1 is then satisfied.

In our description, we do not take into account simultaneous occurence of events. This can be provided for, by performing the described assignment after have incorporated the results of the other simultaneously occuring events. The models in Chapter 4 are more complicated, and so we assume that different events cannot occur simultaneously, in order to keep the analysis tractable.

In Chapter 5 we restrict to controlled single server queues without customer impatience (see [17]). Examples of the models that we will analyse, are the $M/M/1$ queue with a linear holding cost rate (as a function of the number of customers present) and service cost and the two-competing queues model with Poisson arrivals, exponentially distributed service requirements and a quadratic holding cost rate.

For the first model, the $M/M/1$ queue with service rate control, Lippman [30] has shown optimality of a threshold policy. In his seminal paper, he was the first to realise that the value iteration algorithm, which is a numerical algorithm, can be used to inductively to show structural properties of the $n$-horizon optimal policy for controlled $M/M/1$ queues, $n = 1, \ldots$, and hence for the limit. Interestingly enough, in the same paper [30], Lippman showed that choosing the initial function $v_0$ in the value iteration algorithm identically equal to zero, generates a non-increasing sequence of $n$-horizon threshold optimal policies in the above example of the $M/M/1$ queue with service rate control. As far as we know, this monotonicity property has not been further exploited in the subsequent literature.

Lippman's paper [30] has generated a large literature on structural results of $n$-horizon value functions and $n$-horizon optimal policies. Koole [27] has made a next important contribution to this field by introducing event based dynamic programming, as a systematic tool to show propagation of structural properties through each value iteration step. This approach has been succesfully applied in many papers (see [26], [27], [28], [12], etc) to show threshold optimality in one-dimensional controlled queueing models, and to show optimality of index policies in higher dimensional controlled queueing models, or switching curve optimality in two-dimensional models. The fact that there is a threshold or switching curve optimal policy can be exploited to reduce the computational effort for effectively determining the optimal policy.

In Chapter 5, we have delved into the problem whether the above discussed monotonicity of $n$-horizon threshold optimal policies can be enforced by choosing the initial function $v_0$ in the value iteration algorithm cleverly. Can one choose it to obtain a non-decreasing sequence of $n$-stage threshold optimal policies? This would allow to both upper and lower bound the infinite horizon threshold optimal policy, and would further increase numerical efficiency. Indeed, the optimal policy would then be known in a large part of the state space already after one value iteration step for two well chosen initial functions (one generating a non-decreasing sequence of $n$-stage threshold optimal policies and the other generating a non-increasing sequence), and can be directly implemented. Another feature of this procedure would be that, when the upper and lower bounded $n$-horizon threshold optimal policy generated by the two initial functions are equal, this policy must be the desired infinite horizon threshold optimal policy!

This idea indeed has turned out to be a feasible approach. As an example, consider the $M/M/1$ queue with service rate control. Using the fact that the value function should be quadratic as a function of the states (since the cost rate is assumed to be linear), in Chapter 5, Section 5.4 we have been able to construct initial functions with the above monotonicity properties. To concretise this, choose the following parameters:

- arrival rate $\lambda = 0.2$;
- service speed divided into two categories, i.e. low service rate $\mu_1 = 0.3$ and high service rate $\mu_2 = 0.5$;
- holding cost per unit time $c_h(x) := x$;
- operating cost at high speed service per unit time is 1.

Take the following initial function

$$v_0(x) = \frac{5}{3}(x^2 - x) + \frac{16}{3}x, \quad x \geq 0.$$

By virtue of the results in Chapter 5, Section 5.4, this function generates both a non-increasing and a non-decreasing sequence of $n$-stage optimal thresholds. Since the 1-stage optimal threshold equals 1, the optimal threshold equals 1, and so it is optimal to always serve at high speed. As a side remark, the above function must be the value function itself. To illustrate the quality of this result, notice that taking the initial function $v_0$ identically equal to 0, yields 1-stage optimal threshold equal to infinity.

The applicability of our approach crucially hinges on whether such monotonic behaviour of $n$-horizon optimal policies can be enforced as well in more complicated controlled queueing models. In Chapter 5, Section 5.5, we take up the two-competing queues model described above.

In this case, the state space is a two-dimensional one: $\{x \mid x = (x_1, x_2), x_i \in \mathbb{N}_0\}$, where $x_i$ denotes the number of type $i$ customers present, $i = 1, 2$. If the holding cost rate is linear as a function of the state, it is known that an index policy is optimal. One type of customers always has priority, and only when these are absent, the other type is served. To be more specific, let the customers types be indexed by 1 and 2, and denote by $c_i, \mu_i$ the holding cost rate and service rate of type customers of type $i$ customers respectively, $i = 1, 2$. Without loss of generality we may assume that $\mu_1 c_1 \geq \mu_2 c_2$. Then the $c\mu$-rule is optimal, i.e. it is always optimal to prioritise type 1 customers over type 2 customers.

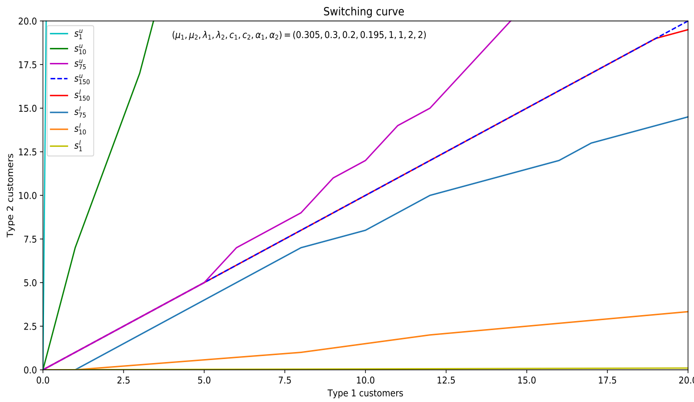


Figure 1.4: $n$-stage optimal upper and lower bounding switching curves after 10, 75, and 150 iterations and the optimal switching curve for $x_1 \leq 19$

In this case, the $n$-horizon optimal policy equals the infinite horizon optimal one, if initial function identically equal to zero is used. This is clearly not a very interesting case to analyse.

Instead, we investigate the case of a quadratic holding cost rate, of the form $c_h(x) = c_1x_1 + \alpha_1x_1^2 + c_2x_2 + \alpha_2x_2^2$, $x_1, x_2 \in \mathbb{N}_0$. First, we show that the $n$-stage optimal policies have a switching curve structure, when the initial function has appropriate monotonicity properties. By a *switching curve structure*, we mean that for any fixed number of customers of one type, the policy has a threshold structure as a function of the other type.

Clearly, the value function must behave as a third power in both state components, because of the fact that the cost rate is a quadratic function. Some cumbersome manipulations turn out to be needed to show that one can choose two initial functions of the required type, such that the $n$-horizon optimal policies form a non-decreasing and a non-increasing sequence of $n$-stage optimal switching curves. As an example, we may take two initial functions in the following generic way

$$v_0^u(x) = \gamma_1^u x_1^3 + \gamma_2^u(x_1+1)^2, \quad v_0^l(x) = \gamma_1^l x_2^3 + \gamma_2^l(x_2+1)^2,$$

with $\gamma_1^i, \gamma_2^i$ given in Lemma 5.5.1, $i = u, l$. The first function yields a non-increasing sequence of $n$-stage optimal switching curves $s_n^u$, $n = 1, \ldots$, and the second a non-decreasing one $s_n^l$, $n = 1, \ldots$.

If we choose (see Example 5.5.3)

$$(\mu_1, \mu_2, \lambda_1, \lambda_2, c_1, c_2, \alpha_1, \alpha_2) = (0.305, 0.3, 0.2, 0.195, 1, 1, 2, 2),$$

then $\gamma_1^u = \gamma_1^l = 6.349$, $\gamma_2^u = 60.09$, and $\gamma_2^l = 59.18$. Convergence to the optimal switching curve for $x_1 \leq 19$ takes place in 145 iterations for the upper bound and in 150 iterations for the lower bound, see Figure 1.4. This picture shows the upper and lower bounding switching curves after 10, 75, and 150 iterations as well. It is noticeable that the optimal switching curve seems to be linear, with gradient $\alpha_1\mu_1/(\alpha_2\mu_2)$. It would be interesting to prove this to be true.

**Open problems**

The following problems would be very interesting to study.

- How can one extend the coupling method from Chapter 4 to allow for control? Can one set up an event based dynamic programming analogon?
- Instead of considering average expected reward MDPs, consider total expected discounted reward MDPs. Then the value function behaves similarly to the immediate rewards as a function of state in the models considered in Chapter 5. This simplifies the construction of suitable initial functions as input of the value iteration algorithm that generate a non-decreasing sequence of lower bounding switching curves and a non-increasing sequence of upper bounding switching curves resepctively. It would be interesting to test this for the models considered.

Do the upper and lower bounding sequence converge quicker than in the average expected reward case?

- In the two-competing queues model with quadratic holding cost, the optimal switching curve appears to be linear for small states. Is it true that it is linear, with the suggested gradient? What about holding cost that behave as a polynomial of degree 3 in the state variable? Are similar properties also true in other two-dimensional control problems, where linear holding costs give rise to an index optimal policy?