Optimal control of admission in service in a queue with impatience and setup costs

Alain Jean-Marie, Emmanuel Hyon

Optimal control of admission in service in a queue with impatience and setup costs

Alain Jean-Marie, Emmanuel Hyon

RESEARCH REPORT

N° 9199

August 2018

Project-Team Maestro

ISSN 0249-6399 ISRN INRIA/RR--9199--FR+ENG

Optimal control of admission in service in a queue with impatience and setup costs

Alain Jean-Marie*, Emmanuel Hyon†

Project-Team Maestro

Research Report n° 9199 — August 2018 — 44 pages

Abstract: We consider a single server queue in continuous time, in which customers must be served before some limit sojourn time of exponential distribution. A customer who is not served before this limit leaves the system: it is impatient. The fact of serving customers and the fact of losing them due to impatience induce costs. The fact of holding them in the queue also induces a constant cost per customer and per unit time. The purpose is to decide when to serve the customers so as to minimize costs. We use a Markov Decision Process with infinite horizon and discounted cost. Since the standard uniformization approach is not applicable here, we introduce a family of approximated uniformizable models, for which we establish the structural properties of the stochastic dynamic programming operator, and we deduce that the optimal policy is of threshold type. The threshold is computed explicitly. We then pass to the limit to show that this threshold policy is also optimal in the original model. A particular care is given to the completeness of the proof. We also illustrate the difficulties involved in the proof with numerical examples.

Key-words: Scheduling, queuing system, impatience, deadline, optimal control, Markov decision processes

* Inria, Univ. Montpellier, LIRMM, CNRS, 840 rue St Priest, F-34395 Montpellier Cedex 05, `Alain.Jean-Marie@inria.fr`.

† Université Paris Nanterre et Sorbonne Universités, LIP6, CNRS, UMR 7606, F-75005, Paris, France, `Emmanuel.Hyon@u-paris10.fr`.

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Contrôle optimal de l'admission en service dans une file d'attente avec impatience et coûts de mise en route

Résumé : Nous considérons un modèle d'une file d'attente à un serveur en temps continu, dans laquelle les clients doivent être servis avant une durée de séjour finie aléatoire, de distribution exponentielle. Un client qui n'est pas servi avant cette limite quitte le système: il est impatient. Le fait de servir les clients et le fait de perdre des clients par impatience induisent des coûts. Le fait de les garder dans la file induit également un coût constant par client et par unité de temps. Il s'agit de décider de façon optimale quand servir les clients. Nous utilisons un processus de décision Markovien à horizon infini et à coûts actualisés. La méthode standard d'uniformisation ne s'appliquant pas à cette situation, nous introduisons une famille de modèles approchés uniformisables pour lesquels nous établissons les propriétés structurelles de l'opérateur de programmation dynamique stochastique, et nous déduisons que la politique optimale est à seuil. Le seuil est calculé explicitement. Nous passons ensuite à la limite pour montrer que cette politique à seuil est également optimale dans le modèle initial. Une attention particulière est apportée à la complétude de la preuve. Nous illustrons également les difficultés rencontrées à l'aide d'exemples numériques.

Mots-clés : Ordonnancement, file d'attente, impatience, échéance, contrôle optimal, processus de décision markovien

1 Introduction

We are interested in this paper in the optimal control of a continuous-time queuing system with impatient customers (or, equivalently said, customers with deadlines). Arrivals follow a Poisson Process and services as well as deadlines follow exponential distributions. The set-up of customer services, the storage of the customers in the queue as well as their departure from the queue due to impatience induce some costs and it has to be decided when to begin the service in order to minimize these costs. At each decision epoch, the controller faces up a trade-off problem between serve/not serve decision and we are interested in this paper to the optimal admission in service policy. We also are motivated to investigate the structural properties satisfied by the optimal policy, and especially the optimality of a special form of structural policies: threshold policies.

Stochastic controlled queuing models, have been largely studied in the literature since their application fields are numerous: networking, resources allocation, inventory control, to quote just a few (see e.g. [10] and references therein). However, most of these works do not consider impatient customers although the phenomenon of impatience, associated with deadlines or "timeouts", has become a major trend in recent years. For non-controlled queues an overview can be found in [14]. Models of controlled queues with impatience appear in several fields of engineering: intrinsincally in real-time systems and yield management, but also in communication networks [18], call centers [12], inventory control [3, 5], big data with volatile data [23]. A variety of optimization problems are solved: admission control in a system [24]; optimal scheduling [6] or optimal routing between queues [20] so as to minimize deadline misses; scheduling in order to minimize long run costs [18]; inventory control in Make-to-order systems [4]; optimal control of the service rates [2]; admission in service in slotted models [15].

From a theoretical point of view, the phenomenon of impatience requires addressing a number of new issues. First, the validity of the Bellman Equation and the existence and the uniqueness of its solution should be proved. Indeed, considering impatience leads to mathematical models with unbounded transition rates. For such models, the standard uniformization method can not be applied and then the usual existence theorems (see Chapter 11 of [22]) can no longer be applied. Recent works have considered this problem and new conditions for the validity of the Bellman Equation has been stated *e.g.* by Guo and Hernandez-Lerma in [13]. Second, when working on the properties of optimal control policies in unbounded models, we are led to approximate them by specific bounded models, to which uniformization can be applied. However the passing of structural results from a bounded model to their generalizations in an unbounded models requires some precise conditions that should be satisfied. Some recent results by Blok and Spieksma ([8] and [9]) give mathematical foundations and the conditions of validity of this approach.

There is a widespread belief (see [12, 17]) that the impatience phenomenon is not compatible with the technique of propagation of structural properties through the dynamic programming operator, and that it can actually destroy structural properties that are known for some queueing systems. For instance, in the call center problem of [12] in which one should schedule two classes of customers, it was quoted that impatience causes the famous $\mu - c$ policy rule to become non optimal. Moreover, the "structure" may actually change, depending on the value of the abandonment rate. For instance, in [2], it is proved that the value function is convex when the abandonment rate is smaller than the service rate; otherwise, the value function seems to be concave according to numerical investigations. A similar situation appears in [18], where the submodularity of the Bellman operator can be proved only when the departure rate is larger than abandonment rate. This means that techniques of proof may or may not succeed, depending on a subtle parametric discussion. The same kind of difficulties occur in our model as it will be seen.

To further stress the difficulties caused by impatience, we can turn to the field of inventory control with perishable items, which has many similarities with queueing models. Indeed, there are connections between inventory models with perishable items and queues with impatience since some inventory models can be expressed as a queueing model with impatience and since the two types of models give rise to fairly close markovian representations. Furthermore, the families of policies considered in the inventory field are (s, S) and (r, Q) policies which are threshold basestock policies. In this field, the structure of the optimal policy in the presence of impatience remains an open question in many problems as stated in [5]. The recent overview on the topic in [16] highlights this point and categorizes continuous-time review problems into three main classes: with no orderings cost (*i.e.* set up cost is null) and positive lead time; with ordering costs and no lead times (*i.e.* service time is zero); with ordering costs and positive lead times. The last case is the most difficult and with less published results. In summary, up to our knowledge, there are only few papers that prove the optimality of a threshold policy on the queue length (or basestock for inventory problems) in queues with impatience. So finding such a proof of optimality, especially in a system featuring both impatience and ordering costs is a noticeable contribution.

Nevertheless, there are papers in which the optimality of threshold policies has been formally shown, sometimes with partial results. For instance, the admission control in a GI/GI/1 queue under heavy traffic with impatience had been studied in [24] but only asymptotic results are obtained: under the heavy traffic assumption the asymptotic (as the arrival rate tends to infinity) optimal policy is of threshold type. Also, [25] proves the optimality of threshold policies, for an inventory model with ordering costs and no lead times. Similarly, [4] considers a Make-to-stock queueing model with impatience when unsatisfied demands are backlogged. In this model, among all the customers in the queue, only backlogged ones are subject to impatience. The control consists of both an admission decision to the system, and an admission decision in service. There are no ordering costs. Threshold policies are shown to be optimal using the propagation of structural properties. The work [7] considers a parametric admission (a customer is admitted with a given probability that must be optimized) in a retrial queue with impatience on retrials and introduces a *Smoothed Rate Truncation* method in order to work with models with bounded transition rates. The model closest to ours but without impatience is the problem of optimally controlling a batch server in a queue. The discrete time model without impatience has been addressed in [21] and has been extended for cases with impatience in [15] (for a batch size of 1 and deterministic service time). The continuous-time model without impatience has been analyzed in [11, 1]. Both papers show that the optimal policy is a threshold policy. A continuous-time model with impatience is studied in [19] which assumes that the batch size is infinite (a clearing of the queue occurs at each service decision) and which also shows that the optimal policy is a threshold policy. But extending the techniques developed in these papers is not straightforward.

Contribution and methodology. In this work we propose a simple queueing model with service control and impatience. This model can be viewed as a generalization of [1, 11] with the introduction of impatience, but restricted to service batches of size one. We prove that an optimal control policy is a threshold policy based on the queue length. The proof is based the construction of a Markov Decision Process and on usual ideas about the propagation of well-chosen structural properties of policies and value functions, via a Bellman operator associated to the problem. However, the implementation of this idea faces here several challenges.

The principal issue is that transition rates of the underlying continuous-time Markov chain are unbounded. The standard reduction of the continuous-time MDP to a discrete-time one using uniformization is not possible. Among the consequences, the "natural" Bellman operator of the problem is not guaranteed to have the required properties. In particular, it is not clear it is contractive and that the Bellman equation has a unique solution, or that the Value Iteration algorithm converges.

In order to handle this situation, we follow the approach recommended by Blok and Spieksma [8] and applied successfully to several similar models in [19, 7]. The idea is to approximate the target model by a sequence of uniformizable models, and use the continuity results established in [8] to deduce properties of the limit model from properties of the approximations. The approximated models are typically truncated models with bounded transition rates, to which the property propagation framework can be applied. As often with truncations, boundary effects may appear that destroy the usually desired properties such as convexity. This effect is countered in [7] by the introduction of a *smoothed* rate truncation. We also use this approach.

Since these technical difficulties are not necessarily explicit in a number of works that deal with impatience in continuous time, we take advantage of this paper to present an in-depth treatment of these technicalities related to unbounded rates. A first step is to adapt the modeling to existing theorems. As it turns out, the strongest mathematical results are presented in a MDP model that does not admit instantaneous costs, whereas our queueing control model does have some. A transformation of models is therefore required. Then it is needed to establish that the solution of the problem is characterized by some Bellman equation. We use results on this question from the work of Guo and Hernandez-Lerma in [13]. Then, it is needed to assess the existence and uniqueness of solutions to this equation, as well as to the ones resulting from approximations. We use results from [8] for this, as well as for obtaining the convergence of the solutions to smoothed and truncated models toward that of the original one. The formal reference to the property propagation framework is taken from [22]. We identify a set of properties that are propagated by the Bellman operator of approximated models and therefore enjoyed by the respective value functions.

The result of the analysis is that, when holding costs are linear, the optimal policy is either to always serve customers, or never serve them. Deciding which of them is optimal is done by comparing the service cost to a combination of the other parameters: loss cost, marginal holding cost, impatience rate and discount rate. Beyond the simplicity of this particular situation, we believe that the techniques presented here will be of a great help to indicate that the propagation of structural properties can also work in more complicated situations. For this purpose, we use a similar approach to both "serve" and "not serve" situations, although results for the second one can be obtained differently.

We also claim that the situation here is not as simple as it looks: we actually demonstrate, through several examples, that simpler, more direct or traditional methods of proof all fail for some reason. This discussion helps understand why each of the technical step is needed.

Organization. This paper is organized as follows: Section 2 deals with the model, while Section 3 study the dynamic control with unbounded rates and Section 4 presents the smoothed rate truncation used here. Section 5 establishes the structural properties propagation while the optimal policy is shown in Section 6 and a methodological discussion is done in part 7. At last, Section 8 concludes this work.

2 Model

This section is devoted to the presentation of the optimization model. In Section 2.1, we describe the features of the queueing model under study, the way it is controlled and the costs that are incurred. In Section 2.2, we explain how to map these requirements to a continuous-time MDP in the formalism of [13] or [8].

2.1 The controlled queueing model

We consider a continuous-time controlled queueing model model, in which customers are assumed to arrive according to a Poisson process with a constant intensity Λ. Once they arrived, customers are stored in an infinite buffer in which they wait for to be admitted in one single server to be processed. Let n be number of customers waiting in the queue (excluding service). The *service admission* decision is made by a controller. Once admitted in service by the controller, the service begins instantly and is not interrupted. The service duration is assumed to follow an exponential distribution of parameter μ.

Customers who wait in the queue are impatient. We assume that customers remain in the queue during a time that has an exponential distribution of parameter α, these durations being independent. In that case, when there are n customers waiting in queue, the rate of impatience is $\alpha(n)$ with $\alpha(n) = n\alpha$. This means that the next departure due to impatience occurs after a duration exponentially distributed with parameter $\alpha(n)$.[1] On the other hand, customers admitted in service are not impatient any more.

In the system we model, costs have several origins:

- starting a service has an instantaneous cost c_B ("B" stands for "batch");
- the departure of a customer due to impatience has an instantaneous cost c_L ("L" stands for "leave");
- costs accumulate continuously over time, at a rate $h(n)$ per time unit that depends on n, the number of customers in the queue.

We wish to find a feedback control policy π that minimizes the infinite-horizon expected discounted cost of policy π, defined as follows. For each initial state x,

$$V^{\pi}(x) \;=\; \mathbb{E}^{\pi}\left(c_B \sum_{k=0}^{\infty} e^{-\theta A_k} + c_L \sum_{k=0}^{\infty} e^{-\theta L_k} + \int_0^{\infty} e^{-\theta t} h(X(t)) \mathrm{d}t \;\middle|\; X(0) = x \right) , \qquad (1)$$

where $\mathbb{E}^{\pi}$ stands for the expectation under the distribution of the process induced by policy π, $\theta > 0$ is the discount parameter, $\{A_k\}_k$ the sequence of times at which an admission in service is made, $\{L_k\}_k$ the sequence of times at which a customer leaves due to impatience, and $\{X(t)\}_t$ is the process representing the number of customers in the waiting room at each instant of time.

2.2 The formal model

We now present the formal mapping of the problem we wish to solve into a continuous-time Markov Decision Process model called *formal model*. This formal model uses more general assumptions than the controlled queueing model requires, and hence contains the controlled queueing problems we consider. It also contains the approximated models which we shall use to handle the difficulty caused by the unbounded transition rates induced by the presence of impatience.

Following [13, 8], a continuous-time MDP model is described by: the state space and the action spaces, the transitions rates $q(y|x,a)$, from state x to y, given that action a is being applied,[2] and the cost $c(x,a)$ that accumulates over time. We proceed with the identification of these elements in our formal model.

[1] Equivalently, the probability of a departure in the next interval of length δt is $\alpha(n)\delta t + o(\delta t)$.

[2] In this model, actions are not taken just at the time of transitions, but continuously over time.

It should be noticed that the model used in [13, 8] involves only running costs, and no instantaneous costs. Therefore we will have to map the instantaneous costs of the controlled queueing model to running costs.

State and action spaces. We denote by $x = (n, b)$ the state of the system. The integer $n \in \mathbb{N}$ denotes the number of customers which are waiting in the queue (excluding service), while the term $b \in \{0, 1\}$ records the status of the server: 0 when the server is idle and 1 when the server is in use. The state space is denoted by $\mathcal{X} = \mathbb{N} \times \{0, 1\}$.

The unique control decision is to admit a customer in service. This is possible only if no customer is already in service, and if some customer is waiting. Accordingly, $\mathcal{A}_x$, the set of controls available in state x, is $\mathcal{A}_x = \{0, 1\}$ for $x = (n, 0)$, $n \in \mathbb{N} \setminus \{0\}$, and $\mathcal{A}_x = \{0\}$ for $x = (0, 0)$ and $x = (n, 1)$, $n \in \mathbb{N}$.

System dynamics. For the sake of generality we assume that the Poisson process of arrivals has a queue-length-dependent intensity. When the number of customers waiting in the queue is n, the arrival intensity is then $\lambda(n)$. We equally assume that while there are n customers waiting in queue, the rate of impatience is $\alpha(n)$, a general function of n. Then we have: when $x = (n, 0)$ and $a = 0$ (no admission in service in an idle server):

$$q(y|x, a) = \begin{cases} \lambda(n) & \text{for} \quad y = (n + 1, 0) \\ \alpha(n) & \text{for} \quad y = (n - 1, 0) \ . \end{cases}$$

When $x = (n, 0)$ and $a = 1$ (admission in service in an idle server), the actual number of customers in the queue is $n - 1$, so that:

$$q(y|x, a) = \begin{cases} \lambda(n-1) & \text{for} \quad y = (n, 1) \\ \alpha(n-1) & \text{for} \quad y = (n - 2, 1) \\ \mu & \text{for} \quad y = (n - 1, 0) \ , \end{cases}$$

provided the state y is in $\mathcal{X}$, and 0 otherwise. Finally, when $x = (n, 1)$ (busy server), then necessarily $a = 0$. The actual number of customers in the queue is n, and:

$$q(y|x, 0) = \begin{cases} \lambda(n) & \text{for } y = (n + 1, 1) \\ \alpha(n) & \text{for } y = (n - 1, 1) \\ \mu & \text{for } y = (n, 0) \, . \end{cases}$$

In the analysis, the total rate of events is useful. Let us denote with $q(x, a)$ the total rate from state x when action a is applied. Given the definitions above, we have:

$$q(x, a) = \begin{cases} \lambda(n) + \alpha(n) + \mu & \text{when } x = (n, 1), n \geq 0 \\ \lambda(n-1) + \alpha(n-1) + \mu & \text{when } x = (n, 0), n \geq 1 \text{ and } a = 1 \\ \lambda(n) + \alpha(n) & \text{when } x = (n, 0), n \geq 0 \text{ and } a = 0 \, . \end{cases}$$

Cost model. According to the specification of the controlled queueing model, costs accumulate in state $x = (n, b)$ at a rate $h(n)$. However, as already mentioned, instantaneous costs do not fit directly in the MDP model. We therefore derive an "equivalent running cost". In order to do this, we switch temporarily to the modeling framework of [22].

Consider the situation where at time T_0 the state is $x = (n, b) \in \mathcal{X}$, action $a \in \mathcal{A}_x$ is taken, and the next event occurs at time T_1. The instantaneous cost is:

$$C_I(x, a) \;=\; c_B \mathbb{1}_{\{(a=1)\cap(n>0)\cap(b=0)\}} . \tag{2}$$

Just after the decision, the number of waiting customers is $n-a$. Therefore, the running cost $h(n-a)$ applies in the interval (T_0, T_1). The duration $T_1 - T_0$ is exponentially distributed with rate $q(x, a)$ and the probability that the event at time T_1 is a customer leaving due to impatience is $\alpha(n-a)/q(x, a)$.

Then, following [22, Chap. 11],[3] the current-value cost incurred between the two events can be expressed as,

$$C(x,a) = \mathbb{E}_x \left(C_I(x,a) + e^{-\theta(T_1-T_0)} c_L \mathbb{1}_{\{\text{impatience at } T_1\}} + \int_{T_0}^{T_1} e^{-\theta(t-T_0)} h(n-a) \mathrm{d}t \,\middle|\, x_0 = x \right) ,$$

and then

$$\begin{aligned} C(x,a) &= C_I(x,a) + c_L \, \frac{\alpha(n-a)}{q(x,a)} \, \mathbb{E}_x \left(e^{-\theta(T_1-T_0)} \right) + h(n-a) \mathbb{E}_x \left(\int_{T_0}^{T_1} e^{-\theta(t-T_0)} dt \right) \\ &= C_I(x,a) + c_L \, \frac{\alpha(n-a)}{q(x,a)} \, \frac{q(x,a)}{q(x,a)+\theta} + h(n-a) \left(\frac{1}{\theta} - \frac{1}{\theta} \mathbb{E}_x \left(e^{-\theta(T_1-T_0)} \right) \right) \\ &= C_I(x,a) + c_L \, \frac{\alpha(n-a)}{q(x,a)+\theta} + \frac{h(n-a)}{q(x,a)+\theta} \, . \end{aligned}$$

If our cost model had only running costs with rate $c(x, a)$, this event-to-event cost would evaluate to $c(x, a)/(q(x, a) + \theta)$. Therefore, for the modeling framework of [13, 8] to correspond to our cost structure, we choose as pseudo-running cost the function:

$$c(x, a) \;=\; (q(x, a) + \theta) C_I(x, a) + c_L \alpha(n-a) + h(n-a) \, .$$

Define the function k as:

$$k(n) \;=\; c_L \alpha(n) + h(n) \, . \tag{3}$$

We summarize this description of the model with the following definition.

Definition 1. *The formal model is the MDP model with transition rates specified in this paragraph, and cost rates given by the function*

$$c(x, a) \;=\; (q(x, a) + \theta) C_I(x, a) + k(n-a)$$

where $C_I(.)$ is defined in (2). Its parameters are c_B, μ, $k(.)$ and the rate functions $\alpha(.)$ and $\lambda(.)$. All these parameters are assumed to be nonnegative.

3 Optimization via stochastic dynamic programming

In this section, we state the optimization problem for the formal model of Definition 1 and characterize its solution. In Propositions 1 and 2 we state that the value function is the unique solution of some Bellman equation. We state the main result of the paper as Theorem 1. The next sections will be devoted to the proof of this theorem.

[3] Some authors advocate the use of Dynkin's formula here, see *e.g.* [19].

The central observation is that our formal model defines transition rates that are not *a priori* unbounded. It is then not possible to use the "classical" framework as described in, say, Puterman [22], which involves *uniformization* to reduce the problem to a discrete-time MDP and then to assert the validity of the Bellman equation and the uniqueness of its fixed point. We apply recent advances in the theory of non-uniformizable MDPs (see [13, 8]) to obtain these results.

3.1 Policies and optimization criterion

In the context of unbounded-rate MDPs, one concentrates on the set of *Stationary Markov Deterministic Policies*. Those policies are characterized by a single, deterministic decision rule which maps the current state to an action and which is applied continuously over time. We denote with F the set of such policies.[4]

For a given policy π, let $V^\pi(x)$ be the value function, representing the average total discounted cost of policy π when the initial state of the process is x, defined as:

$$V^\pi(x) = \mathbb{E}\left(\int_0^\infty e^{-\theta t} c(X(t), \pi(t))\, \mathrm{d}t \;\middle|\; x_0 = x \right) . \tag{4}$$

Here, $\pi(t)$ denotes the action prevailing at time t according to the policy π, given that the decisions prescribed by this policy are applied at the instants where *events* occur, and $\theta > 0$ is a discount factor.

The general objective of optimal control is to find, for every x, a policy π that minimizes the criterion $V^\pi(x)$. Classical results (*e.g.* [22]) on discounted, infinite-horizon, time-homogeneous Markovian optimal control have shown that, in many cases, policies from set F are globally optimal. However, for continuous-time models that are not uniformizable, which is the case here, this certainty is not known. It is then a practically reasonable objective to consider the optimization problem: find

$$\pi^* = \arg\min_{\pi \in F} V^\pi(x) ,$$

provided this "min" is attained.

3.2 Bellman Equations

The first result is that the optimal value function satisfies a dynamic programming equation (Discounted Cost Optimality Equation, DCOE), generally known as a Bellman equation.

Proposition 1. *In the formal model, assume that the function $\lambda(\cdot)$ is bounded, $\alpha(\cdot)$ and $k(\cdot)$ are bounded by polynomials. Then the value function of the problem satisfies the DCOE:*

$$\begin{aligned} V(n,0) \;\; = \;\; \min\Big\{ c_B + \frac{1}{\lambda(n-1)+\alpha(n-1)+\mu+\theta} \big[& k(n-1) + \lambda(n-1)V(n,1) \\ & + \alpha(n-1)V(n-2,1) + \mu V(n-1,0) \big], \end{aligned} \tag{5}$$

$$\frac{1}{\lambda(n)+\alpha(n)+\theta} \left[k(n) + \lambda(n)V(n+1,0) + \alpha(n)V(n-1,0) \right] \Big\} \tag{6}$$

for $n \geq 1$,

$$V(0,0) \;\; = \;\; \frac{1}{\lambda(0)+\theta} \left[k(0) + \lambda(0)V(1,0) \right], \tag{7}$$

[4] This is the notation from [13]; this set is denoted as $\mathcal{D}$ in [8].

$$V(n,1) = \frac{1}{\lambda(n)+\alpha(n)+\mu+\theta}\left[k(n)+\lambda(n)V(n+1,1)+\alpha(n)V(n-1,1)+\mu V(n,0)\right] , \tag{8}$$

for $n \geq 0$.

Moreover, the Bellman equations have a unique solution and provide an optimal feedback control:

Proposition 2. *Under the assumptions of Proposition 1, the DCOE equations (5)–(8) have a unique solution. In addition, any function* $\gamma : \mathcal{X} \to \{0,1\}$ *which realizes the* "min" *in (5)–(6) is optimal in* F.

The proofs of these results are in Appendix A and Appendix B, respectively.

3.3 Optimal policy

Our objective for the present paper is to solve the optimization problem described in Section 2.1 by finding explicitly the optimal control in the following special case, called the "*base model*".

Definition 2 (Base Model)**.** *The base model is the model defined in Section 2.2 with:*

$$\lambda(n) = \Lambda, \qquad \alpha(n) = n\alpha, \qquad k(n) = nc_Q.$$

This assumption on $k(.)$ corresponds to the assumption that $h(n)$ is linear in the controlled queueing model of Section 2.1. Indeed, if $h(n) = nc_H$ then given the definition (3), $k(n) = n(\alpha c_L + c_H)$ is also linear. The correspondence holds with $c_Q = \alpha c_L + c_H$.

The optimal control follows policies of special form. Define the following policies in F:

Definition 3 ("No Service" and "Always Serve" policies)**.** π_{NS} *("no service") is the policy in* F *that selects* $a = 0$ *in every state, i.e.* $\pi_{NS}(x) = 0$, $\forall x \in \mathcal{X}$.
π_{AS} *("always serve") is the policy in* F *that selects* $a = 1$ *in every relevant state, i.e.* $\pi_{AS}(x) = 1$, $\forall x = (n, 0)$, $n \geq 1$.

The principal result of the paper states as:

Theorem 1. *Consider the Base Model. Then:*

a) if $c_Q < c_B(\alpha + \theta)$*, then* π_{NS} *is optimal;*

b) if $c_Q > c_B(\alpha + \theta)$*, then* π_{AS} *is optimal;*

c) if $c_Q = c_B(\alpha + \theta)$*, then both* π_{AS} *and* π_{NS} *are optimal.*

The proof of this theorem will be given in Section 6.2. In the next sections, we introduce the concepts needed in the proof: approximated uniformizable models in Section 4 and their structural properties in Section 5.

4 Smoothed and uniformized models

The main topic of this section is the presentation of approximations of the formal model that are uniformizable. The usual way of defining such approximations involves truncation. As shown by Bhulai, Brooms and Spieksma in [7], the addition of *smoothing* may endow the approximations with more interesting structural properties. We therefore apply their "Smoothed Rate Truncation" (SRT) technique. We introduce in Section 4.1 several sets of assumptions on the model data, then we define in Section 4.2 the transition operators which we will later study in Section 5.

4.1 Assumptions

Consider the following assumptions bearing on the formal model of Section 2.2.

Assumption 1 (approximation)**.** *There exists an integer number $N \geq 1$, such that:*

a) The function $\alpha(\cdot) : \mathbb{N} \to \mathbb{R}_+$ is given by

$$\alpha(n) \;=\; \min(n, N)\,\alpha\,.$$

b) The function $\lambda(\cdot) : \mathbb{N} \to \mathbb{R}_+$ is given by

$$\lambda(n) \;=\; \Lambda\,\left(1 - \frac{n}{N}\right)^{+}. \tag{9}$$

c) The function $k(\cdot) : \mathbb{N} \to \mathbb{R}_+$ is given by

$$k(n) \;=\; \min(n, N+1)c_Q.$$

We stress the fact that the uniformized models are used as a mathematical device for obtaining results on the Base Model. As such, we are not interested in their “physical” relevance. In particular, part c) of this assumption does not correspond to the actual costs of the model, as they were identified in Section 2 and Definition 2. What is important is that when N goes to infinity, this model approaches the Base Model, in a sense to be discussed later on.

4.2 Definition of operators

We introduce now two operators $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$ associated with the specific policies π_{AS} and π_{NS} introduced in Definition 3, as well as the dynamic programming operator $T^{(u)}$ of the problem. If Assumption 1 holds, it is possible to define $\tilde{\Lambda} := \Lambda + N\alpha + \mu$, $\Lambda(n) = \tilde{\Lambda} - \alpha(n) - \lambda(n)$ and for $n \geq 1$,

$$c(n) \;:=\; c_B\,\frac{\lambda(n-1) + \alpha(n-1) + \mu + \theta}{\tilde{\Lambda} + \theta}\,. \tag{10}$$

The rate $\tilde{\Lambda}$ is the maximal transition rate in the model and $\Lambda(n)$ is the rate of “dummy transitions” occurring in state n as a result of uniformization.

Let $\mathcal{V}$ be the space of all functions from $\mathcal{X}$ to $\mathbb{R}$. Define the following operators mapping $\mathcal{V}$ to $\mathcal{V}$:

Definition 4 (No-Service operator, smoothed and uniformized model)**.** *Suppose that Assumption 1 holds. Let $\tilde{\Lambda} := \Lambda + N\alpha + \mu$, $\Lambda(n) = \tilde{\Lambda} - \alpha(n) - \lambda(n)$.*

The “no service” operator for the smoothed and uniformized model, $T^{(u)}_{NS}$, is defined as, for any $V \in \mathcal{V}$:

$$(T^{(u)}_{NS}V)(n,0) \;=\; \frac{1}{\tilde{\Lambda}+\theta}\left[k(n) + \alpha(n)V(n-1,0) + \lambda(n)V(n+1,0) + \Lambda(n)V(n,0)\right], \tag{11}$$

$$\begin{aligned}(T^{(u)}_{NS}V)(n,1) \;=\; & \frac{1}{\tilde{\Lambda}+\theta}\big[k(n) + \alpha(n)V(n-1,1) + \lambda(n)V(n+1,1) \\ & + \mu V(n,0) + (\Lambda(n)-\mu)V(n,1)\big]\end{aligned} \tag{12}$$

for $n \geq 0$.

Definition 5 (Always-Service operator, smoothed and uniformized model)**.** *Suppose that Assumption 1 holds. Let* $\tilde{\Lambda} := \Lambda + N\alpha + \mu$, $\Lambda(n) = \tilde{\Lambda} - \alpha(n) - \lambda(n)$ *and let* $c(n)$ *be defined as in (10).*

The "always service" operator for the smoothed and uniformized model, $T^{(u)}_{AS}$*, is defined as, for any* $V \in \mathcal{V}$*:*

$$\begin{aligned}
(T^{(u)}_{AS}V)(n,0) &= c(n) + \frac{1}{\tilde{\Lambda}+\theta}\Big[k(n-1) + \alpha(n-1)V(n-2,1) + \lambda(n-1)V(n,1) \\
&\qquad + \mu V(n-1,0) + (\Lambda(n-1)-\mu)V(n,0)\Big] \qquad (13)\\
(T^{(u)}_{AS}V)(0,0) &= \frac{1}{\tilde{\Lambda}+\theta}\left[k(0) + \lambda(0)V(1,0) + \Lambda(0)V(0,0)\right] \qquad (14)\\
(T^{(u)}_{AS}V)(n,1) &= \frac{1}{\tilde{\Lambda}+\theta}\big[k(n) + \alpha(n)V(n-1,1) + \lambda(n)V(n+1,1) \\
&\qquad + \mu V(n,0) + (\Lambda(n)-\mu)V(n,1)\big]\ , \qquad (15)
\end{aligned}$$

for $n \geq 1$ *in (13) and* $n \geq 0$ *in (15).*

Observe that when $n = 0$ in (12) or (15), the undefined value $V(-1,1)$ is always multiplied by $\alpha(0) = 0$. Observe also that $(T^{(u)}_{AS}V)(0,0) = (T^{(u)}_{NS}V)(0,0)$ and $(T^{(u)}_{AS}V)(n,1) = (T^{(u)}_{NS}V)(n,1)$ for all $n \geq 1$.

Definition 6. *Assume Assumption 1 holds. Define the operator* $T^{(u)}$ *as, for any* $V \in \mathcal{V}$*:*

$$T^{(u)}V = \min\{T^{(u)}_{AS}V, T^{(u)}_{NS}V\}.$$

The operator $T^{(u)}$ is the dynamic programming operator of the uniformized approximated model. In the sequel, we shall use indifferently the notations $(T^{(u)}V)(x) = T^{(u)}V(x)$ for the value of the function obtained by applying $T^{(u)}$ on function V. Likewise for operators $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$.

4.3 Bellman equations

Theorem 2. *Consider the formal control model of Definition 1 with functions* $\lambda(\cdot)$*,* $\alpha(\cdot)$ *and* $k(\cdot)$ *satisfying Assumption 1.*

The value function $V^{(u)}$ *of this problem is the unique solution of the Bellman equation:*

$$T^{(u)}V \;=\; V\ . \qquad (16)$$

In addition, any function γ *such that* $\gamma(x) = \arg\min\{T^{(u)}_{AS}V^{(u)}(x), T^{(u)}_{NS}V^{(u)}(x)\}$ *is optimal in* F*.*

Proof. The statement is a corollary of Proposition 1 and Proposition 2, given the definitions of $T^{(u)}_{AS}$, $T^{(u)}_{NS}$ and $T^{(u)}$ in Definitions 4– 6. These propositions indeed apply since under Assumption 1, and for every fixed N, $\lambda(\cdot)$, $\alpha(\cdot)$ and $k(\cdot)$ (hence $h(\cdot)$) are bounded. □

5 Structural properties of smoothed and uniformized models

We now turn to the analysis of structural properties for the uniformized models defined in 4. We define a set of structural properties of value functions in Section 5.1, and we show that these properties propagate through the operators $T^{(u)}_{NS}$ and $T^{(u)}_{AS}$ in 5.2. These results will be exploited in Section 6.

For functions $V \in \mathcal{V}$, define the following functional: $\Delta_n : \mathcal{V} \to \mathcal{V}$ as:

$$(\Delta_n V)(n,b) \;=\; V(n+1,b) - V(n,b), \qquad (n,b) \in \mathcal{X}.$$

5.1 Definition of properties

One defines here the following properties for functions $V \in \mathcal{V}$.

First of all, we should make it precise that, when we write that a function $f(n)$ defined from $\mathbb{N} \to \mathbb{R}$ is increasing over a range $n \in [a_1..a_2]$, or, more commonly, for $a_1 \le n \le a_2$, we mean that $f(n+1) \ge f(n)$ for $a_1 \le n < a_2$. Secondly, when we write that a function $f(n)$ defined from $\mathbb{N} \to \mathbb{R}$ is convex over a range $n \in [a_1..a_2]$, or, more commonly, for $a_1 \le n \le a_2$, we mean that $f(n+2) - 2f(n+1) + f(n) \ge 0$ for $a_1 \le n < a_2 - 1$. If $a_1 \le a_2 \le a_1 + 1$, this requirement is void.

In all situations, writing that a function f has some property over the range $[a_1..a_2]$ involves the values $f(a_1), \ldots, f(a_2)$ and no other value.

Definition 7 (Properties of value functions)**.** *Let $V \in \mathcal{V}$. Assume Assumption 1 holds for some N. Let $\Phi := N\alpha$. We say that V has property:*

Picx *if $n \mapsto \Delta_n V(n,0)$ is positive and increasing for $0 \le n \le N$;*

P1 *if*

a) $\Delta_n V(n,0) \le c_Q/(\alpha + \theta + \Lambda/N)$ *for all* $0 \le n \le N$*;*

b) $\Delta_n V(0,0) \ge c_B$

P2 *if $n \mapsto (\Phi - \alpha(n))\Delta_n V(n,0) + \alpha(n-1)\Delta_n V(n-1,0)$ is increasing for $1 \le n \le N$;*

P3 *if $n \mapsto \lambda(n)\Delta_n V(n+1,0) + (\Lambda - \lambda(n-1))\Delta_n V(n,0)$ is increasing for $1 \le n \le N$;*

P4 *if $V(n+1,0) = V(n,1) + c_B$, for all $0 \le n \le N$;*

Plin *if V satisfies:*

$$V(n,0) = V(n,1) = \frac{c_Q}{\alpha + \theta + \Lambda/N}\left(n + \frac{\Lambda}{\theta}\right) \,,\; \forall\, 0 \le n \le N\,. \tag{17}$$

In the remainder of this section, we will show results grouped in three categories:

1. simple implications between the properties of Definition 7, in Section 5.2
2. propagation of properties by operator $T^{(u)}_{AS}$, independently of the fact that it realizes optimality in the dynamic programming equation, in Section 5.4,
3. propagation of properties by operator $T^{(u)}_{NS}$, independently of the fact that it realizes optimality in the dynamic programming equation, in Section 5.5.

5.2 Implications and identities for properties

We show now that P2 and P3 are a consequence of Picx.

Lemma 1. *Assume that Assumption 1 holds, and let $V \in \mathcal{V}$ satisfy Picx. Then V satisfies P2 and P3.*

The proof uses the following lemma. This lemma is stated in a quite generic way: although all its statements are not used in the remainder, it is useful to realize what features of the functions $\lambda(\cdot)$ and $\alpha(\cdot)$ are essential in the analysis.

Lemma 2. *Let f: $\mathbb{N} \to \mathbb{R}$ be a function which is positive and increasing for $a_1 \leq n \leq a_2$. Let g: $\mathbb{N} \to \mathbb{R}$ be a positive and convex function, bounded above by G. Let h: $\mathbb{N} \setminus \{0\} \to \mathbb{R}$ be the function defined by*

$$h(n) \ := \ g(n)f(n) \ + \ (G - g(n-1))f(n-1) \ .$$

Then h is positive and:

a) h is increasing for $a_1 + 1 \leq n \leq a_2$.

b) If in addition $a_1 \geq 1$ and $g(a_1 - 1) = G$, h is increasing for $a_1 \leq n \leq a_2$.

c) If in addition $g(a_2 + 1) = 0$, h is increasing for $a_1 + 1 \leq n \leq a_2 + 1$.

Proof. For $n \geq 1$, we have:

$$\begin{aligned}
h(n+1) - h(n) &= g(n+1)f(n+1) \ + \ (G - g(n))f(n) \\
&\quad - [g(n)f(n) \ + \ (G - g(n-1))f(n-1)] \\
&= g(n+1)(f(n+1) - f(n)) \\
&\quad +(g(n+1) + G - g(n) - g(n) - G + g(n-1))f(n) \\
&\quad +(G - g(n-1))(f(n) - f(n-1)) \\
&= g(n+1)\ (f(n+1) - f(n)) \qquad (18) \\
&\quad +(g(n+1) - 2g(n) + g(n-1))\ f(n) \\
&\quad +(G - g(n-1))\ (f(n) - f(n-1)) \ . \qquad (19)
\end{aligned}$$

Since f is increasing on the range $a_1 \leq n \leq a_2$ and g positive, then the first of these three terms is positive for $a_1 \leq n \leq a_2 - 1$ The second one is positive for $a_1 + 1 \leq n \leq a_2$ (or for $a_1 \leq n \leq a_2$ when $a_1 \geq 1$). Indeed, f is positive and g is convex for any $n \in \mathbb{N}$. The third one is positive for $a_1 + 1 \leq n \leq a_2$ since f is increasing on the range $a_1 \leq n \leq a_2$, and $G - g(n-1)$ is positive. The increments of h are positive for $a_1 + 1 \leq n \leq a_2 - 1$ so the part a) of the lemma is proved.

If $a_1 \geq 1$ and $g(a_1 - 1) = G$, then (19) vanishes when $n = a_1$, so the sum remains positive even for $n = a_1$. This proves b).

If $g(a_2 + 1) = 0$, then (18) vanishes when $n = a_2$, so the sum remains positive even for $n = a_2$. This proves c). □

Proof of Lemma 1. For Property P2, let $f(n) = \Delta_n V(n, 0)$. This is a positive function, increasing for $0 \leq n \leq N$ since V satisfies Picx. Let also $g(n) = \Phi - \alpha(n)$. This function is: a) positive and convex; b) bounded by $\Phi = N\alpha$. Lemma 2 a) applies therefore with $a_1 = 0$, $a_2 = N$ and $G = \Phi$, to conclude that $n \mapsto (\Phi - \alpha(n))\Delta_n V(n, 0) + \alpha(n-1)\Delta_n V(n-1, 0)$ is increasing for $1 \leq n \leq N$; in other words, P2 holds.

For Property P3, let $f(n) = \Delta_n V(n+1,0)$. This is a positive function, increasing for $0 \leq n \leq N-1$. Let also $g(n) = \lambda(n)$. This function is: a) positive and convex; b) bounded by Λ; c) such that $g(N) = 0$. Lemma 2 a) and c) apply therefore with $a_1 = 0$, $a_2 = N-1$ and $G = \Lambda$, to conclude that $n \mapsto \lambda(n)\Delta_n V(n+1,0) + (\Lambda - \lambda(n-1))\Delta_n V(n,0)$ is increasing for $1 \leq n \leq N$; in other words, P3 holds. □

5.3 Identities for difference operators

In this part we give the expressions of Δ_n for the operators $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$.

Observe that in the case where P4 holds, the value $(T^{(u)}_{AS} V)(n,0)$ defined in (13) can be rewritten eliminating terms $V(m,1)$. We formulate this, and other consequences, in the next lemmas.

Lemma 3 (Formula of $\Delta_n T^{(u)}_{AS} V$)**.** *Let Assumption 1 hold. Assume V satisfies P4. Then for all $1 \leq n \leq N$,*

$$\begin{aligned}(\tilde{\Lambda}+\theta)T^{(u)}_{AS}V(n,0) \;=\;& c_B\,(\mu+\theta) + (n-1)c_Q \\ &+\alpha(n-1)V(n-1,0) + \lambda(n-1)V(n+1,0) \\ &+\mu V(n-1,0) + (\Lambda(n-1)-\mu)\,V(n,0)\,. \end{aligned} \tag{20}$$

As a consequence, we have, for $\Delta_n T^{(u)}_{AS} V(n,0) := (T^{(u)}_{AS}V)(n+1,0) - (T^{(u)}_{AS}V)(n,0)$

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_n T^{(u)}_{AS}V(n,0) \;=\;& c_Q + \mu\Delta_n V(n-1,0) \\ &+\lambda(n)\Delta_n V(n+1,0) + \big(\Lambda-\lambda(n-1)\big)\Delta_n V(n,0) \\ &+\big(\Phi-\alpha(n)\big)\Delta_n V(n,0) + \alpha(n-1)\Delta_n V(n-1,0)\end{aligned} \tag{21}$$

for $1 \leq n < N$, and

$$(\tilde{\Lambda}+\theta)\Delta_n T^{(u)}_{AS}V(0,0) \;=\; c_B\,(\mu+\theta) + \Lambda\Delta_n V(1,0) + \Phi\Delta_n V(0,0)\,. \tag{22}$$

Proof. In Formula (13), we see that the only instances of $V(m,1)$ are $V(n,1)$ and $V(n-2,1)$. Since by P4, $V(m,1) = V(m+1,0) - c_B$ for $0 \leq m \leq N$, we can therefore eliminate them and obtain, for all $1 \leq n \leq N$:[5]

$$\begin{aligned}(\tilde{\Lambda}+\theta)T^{(u)}_{AS}V(n,0) \;=\;& c_B\,(\lambda(n-1)+\alpha(n-1)+\mu+\theta) \\ &+k(n-1) + \alpha(n-1)(V(n-1,0)-c_B) \\ &+\lambda(n-1)(V(n+1,0)-c_B) + \mu V(n-1,0) + (\Lambda(n-1)-\mu)V(n,0)\end{aligned}$$

which is rearranged into (20) remembering that $k(n-1) = (n-1)c_Q$.

Next, we express the difference $\Delta_n T^{(u)}_{AS}V(n,0) := (T^{(u)}_{AS}V)(n+1,0) - (T^{(u)}_{AS}V)(n,0)$, for $1 \leq n < N$. We use (20) and obtain:

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_n T^{(u)}_{AS}V(n,0) \;=\;& c_Q + \alpha(n)V(n,0) + \lambda(n)V(n+2,0) \\ &+\mu V(n,0) + (\Lambda(n)-\mu)V(n+1,0) \\ &-[\alpha(n-1)V(n-1,0) + \lambda(n-1)V(n+1,0) \\ &\quad + \mu V(n-1,0) + (\Lambda(n-1)-\mu)V(n,0)]\,.\end{aligned}$$

[5] The value of $V(n-2,0)$ in (13) is not defined when $n=1$ but is then multiplied by $\alpha(0)=0$ in the formulas.

We replace $\Lambda(n)$ by its value $\Lambda + \Phi + \mu - \alpha(n) - \lambda(n)$ to get for all $1 \leq n < N$:

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_n T_{AS}^{(u)} V(n,0) &= c_Q + \alpha(n)V(n,0) + \lambda(n)V(n+2,0) \\ &\quad + \mu V(n,0) + [\Lambda + \Phi - \alpha(n) - \lambda(n)]V(n+1,0) \\ &\quad - \big[\alpha(n-1)V(n-1,0) + \lambda(n-1)V(n+1,0) \\ &\qquad + \mu V(n-1,0) + [\Lambda + \Phi - \alpha(n-1) - \lambda(n-1)]V(n,0)\big].\end{aligned}$$

Finally, grouping the terms depending on μ, λ and α leads to (21).

In the case $n = 0$, we have, from (20) for $n = 1$ (observe that $N \geq 1$ due to Assumption 1) and (14),

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_n T_{AS}^{(u)} V(0,0) &= (\tilde{\Lambda}+\theta)(T_{AS}^{(u)} V(1,0) - T_{AS}^{(u)} V(0,0)) \\ &= c_B\ (\mu+\theta) \\ &\quad + \alpha(0)V(0,0) + \lambda(0)V(2,0) + \mu V(0,0) + (\Lambda(0) - \mu)V(1,0) \\ &\quad - [k(0) + \lambda(0)V(1,0) + \Lambda(0)V(0,0)] \\ &= c_B\ (\mu+\theta) + \Lambda\,[V(2,0) - V(1,0)] + \Phi\,[V(1,0) - V(0,0)]\ .\end{aligned}$$

In the last step of this derivation, we have used the fact that $\alpha(0) = 0$ and $\lambda(0) = \Lambda$, from Assumption 1, $k(0) = 0$ and the value $\Lambda(0) = \Phi + \mu$. This expression simplifies into (22). □

We define the difference $\Delta_q T^{(u)} V$ which represents the difference between the choice of serving a customer or not serving a customer. For functions $V \in \mathcal{V}$ and for all $n \geq 0$, define the following difference $\Delta_q T^{(u)} : \mathcal{V} \to \mathbb{R}$ as:

$$\Delta_q T^{(u)} V(n) := (T_{AS}^{(u)} V)(n,0) - (T_{NS}^{(u)} V)(n,0). \tag{23}$$

Properties of this operator are summarized in the following result.

Lemma 4 (Properties of $\Delta_q T^{(u)} V$)**.** *Assume that Assumption 1 holds. The function $\Delta_q T^{(u)} V$ defined in (23) has the following properties:*

i) $\Delta_q T^{(u)} V(0) = 0$ *for every* $V \in \mathcal{V}$*;*

ii) If V *satisfies P4, then for* $1 \leq n \leq N$*:*

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_q T^{(u)} V(n) &= (\mu+\theta)\ (c_B - \Delta_n V(n-1,0)) - c_Q \\ &\quad + (\alpha+\theta)\ \Delta_n V(n-1,0) + (\Lambda/N)\ \Delta_n V(n,0)\ .\end{aligned} \tag{24}$$

iii) If V *satisfies P4, then for* $n = N+1$*,* $\Delta_q T^{(u)} V(n)$ *is given by:*

$$(\tilde{\Lambda}+\theta)\Delta_q T^{(u)} V(N+1) = (\mu+\theta)c_B - c_Q - \mu \Delta_n V(N,0)\ . \tag{25}$$

iv) If V *satisfies Plin, then for* $1 \leq n \leq N$*,*

$$(\tilde{\Lambda}+\theta)\Delta_q T^{(u)} V(n) = \left(c_B - \frac{c_Q}{\alpha+\theta+\Lambda/N}\right)(\mu+\theta+\alpha(n-1)+\lambda(n-1))\ . \tag{26}$$

Proof. $i)$ The case $n = 0$ is immediate since $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$ have the same value at $(n, 0)$.

$ii)$ Since V satisfies P4, Lemma 3 applies and we use (20) for $T^{(u)}_{AS}V(n,0)$ and (11) for $T^{(u)}_{NS}V(n,0)$ to obtain, for all $1 \le n \le N$,

$$\begin{aligned}
(\tilde{\Lambda}+\theta)\Delta_q T^{(u)}V(n) &= c_B\,(\mu+\theta) + (n-1)c_Q \\
&\quad +\alpha(n-1)V(n-1,0) + \lambda(n-1)V(n+1,0) \\
&\quad +\mu V(n-1,0) + (\Lambda(n-1)-\mu)V(n,0) \\
&\quad -[nc_Q + \alpha(n)V(n-1,0) + \lambda(n)V(n+1,0) + \Lambda(n)V(n,0)] \\
&= c_B\,(\mu+\theta) - c_Q \\
&\quad +[\alpha(n-1) - \alpha(n) + \mu]\,V(n-1,0) \\
&\quad +[\Lambda(n-1) - \mu - \Lambda(n)]\,V(n,0) \\
&\quad +[\lambda(n-1) - \lambda(n)]\,V(n+1,0) \\
&= c_B\,(\mu+\theta) - c_Q \\
&\quad +[\alpha(n-1) - \alpha(n) + \mu]\,V(n-1,0) \\
&\quad +[-\lambda(n-1) - \mu + \lambda(n) + \alpha(n) - \alpha(n-1)]\,V(n,0) \\
&\quad +[\lambda(n-1) - \lambda(n)]\,V(n+1,0)\,.
\end{aligned}$$

We replaced $\Lambda(n)$ by its value in this last equation. Now we group the terms in $\alpha(n-1) - \alpha(n) + \mu$ and these in $\lambda(n-1) - \lambda(n)$ to identify the instances of Δ_n. Finally we use the specific forms of $\alpha(\cdot)$ and $\lambda(\cdot)$ in Assumption 1. In a last step, we add and subtract the term $(\mu+\theta)\Delta_n V(n-1,0)$ to get:

$$\begin{aligned}
(\tilde{\Lambda}+\theta)\Delta_q T^{(u)}V(n) &= c_B\,(\mu+\theta) - c_Q \\
&\quad +(\alpha-\mu)\,(\Delta_n V)(n-1,0) \\
&\quad +(\Lambda/N)\,(\Delta_n V)(n,0)\,, \qquad (27)\\
&= c_B\,(\mu+\theta) - c_Q - (\mu+\theta)\Delta_n V(n-1,0) \\
&\quad +(\alpha-\mu)\,\Delta_n V(n-1,0) + (\mu+\theta)\Delta_n V(n-1,0) \\
&\quad +(\Lambda/N)\,\Delta_n V(n,0)\,,
\end{aligned}$$

and then the final form (24).

$iii)$ When $n = N+1$, consider Equation (11) and Equation (13). Since Assumption 1 holds then all the terms in "λ" disappear. Computing the difference, and using the fact that $k(N+1) - k(N) = c_Q$ according to Assumption 1 c), we obtain

$$(\tilde{\Lambda}+\theta)\Delta_q T^{(u)}V(N+1) = c_B(\mu+\theta) - c_Q + \alpha(N)(c_B + V(N-1,1) - V(N,0)) - \mu\Delta_n V(N,0)\,.$$

Since V satisfies P4, $c_B + V(N-1,1) - V(N,0) = 0$ and we obtain Equation (25).

$iv)$ Going back to the definitions (11) and (13), we obtain respectively, for any $n \ge 1$,

$$\begin{aligned}
(\tilde{\Lambda}+\theta)T^{(u)}_{NS}V(n,0) &= k(n) + \tilde{\Lambda}V(n,0) - \alpha(n)\Delta_n V(n-1,0) + \lambda(n)\Delta_n V(n,0) \\
(\tilde{\Lambda}+\theta)T^{(u)}_{AS}V(n,0) &= k(n-1) + \tilde{\Lambda}V(n,0) + \alpha(n-1)(c_B + V(n-2,1) - V(n,0)) \\
&\quad + \lambda(n-1)(c_B + V(n,1) - V(n,0)) - \mu\Delta_n V(n-1,0) + c_B(\mu+\theta).
\end{aligned}$$

If V satisfies Plin, that is, (17), then for all $1 \le n \le N$, $V(n,1) = V(n,0)$ and also, for all $0 \le n < N$,

$$\Delta_n V(n,0) = V(n+1,0) - V(n,1) = \Delta := \frac{c_Q}{\alpha+\theta+\Lambda/N}\,.$$

The occurrences of $V(m,1)$ in $(\tilde{\Lambda}+\theta)T^{(u)}_{AS}V(n,0)$ involve only values of m with $1 \le m \le N$. Also, occurrences of $\Delta_n V(m,0)$ involve only values of m with $1 \le m < N$. Only in $(\tilde{\Lambda}+\theta)T^{(u)}_{NS}V(n,0)$, for $n = N$, does the value of $\Delta_n V(N,0)$ appear. This value is unknown but is multiplied by $\lambda(N) = 0$ so we can ignore it. In other words, the identity

$$\lambda(n)\Delta_n V(n,0) = \lambda(n)\frac{c_Q}{\alpha+\theta+\Lambda/N}$$

holds for all n, $1 \le n \le N$.

Consequently, for $1 \le n \le N$,

$$\begin{aligned}(\tilde{\Lambda}+\theta)\Delta_q T^{(u)}V(n) &= k(n-1) + \alpha(n-1)(c_B - 2\Delta) + \lambda(n-1)c_B - \mu\Delta + c_B(\mu+\theta) \\ &\quad - [k(n) - \alpha(n)\Delta + \lambda(n)\Delta] \\ &= c_B(\mu+\theta+\alpha(n-1)+\lambda(n-1)) - c_Q - \Delta[\mu+\lambda(n)+2\alpha(n-1)-\alpha(n)] \\ &= c_B(\mu+\theta+\alpha(n-1)+\lambda(n-1)) - \Delta(\mu+\theta+\lambda(n-1)+\alpha(n-1)) \\ &\quad + \Delta(\alpha+\theta+\lambda(n-1)-\lambda(n)) - c_Q\ .\end{aligned}$$

In the last rewriting, we have used the specific properties of $\alpha(n)$ deriving from Assumption 1. The last line cancels out since $\Delta(\alpha+\theta+\Lambda/N) = c_Q$. There remains (26).

□

5.4 Invariant properties for operator $T^{(u)}_{AS}$

Throughout this section, we shall assume that

$$c_B \ \le\ \frac{c_Q}{\alpha+\theta+\Lambda/N}\ . \tag{28}$$

The main result of this section is the following:

Proposition 3 (Propagation for $T^{(u)}_{AS}$)**.** *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies properties Picx, P1 a), P1 b) and P4. Then $T^{(u)}_{AS}V$ also satisfies these four properties.*

We shall decompose the proof in separate lemmas, for each individual propagation.

Lemma 5 (Propagation of P4 for $T^{(u)}_{AS}$)**.** *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies property P4. Then $T^{(u)}_{AS}V$ also satisfies P4.*

Proof. We identify the term $T^{(u)}_{AS}V(n,0)$ evaluated at $n+1$ in (13), and $T^{(u)}_{AS}V(n,1)$ with (15). We find that for $n \ge 0$:

$$T^{(u)}_{AS}V(n+1,0) - T^{(u)}_{AS}V(n,1) \ =\ c(n+1)\ +\ \frac{1}{\tilde{\Lambda}+\theta}\,(\Lambda(n)-\mu)\,[V(n+1,0)-V(n,1)]\ .$$

Since V satisfies P4, $V(n+1,0) = V(n,1) + c_B$ for all $0 \le n \le N$ and it remains:

$$T^{(u)}_{AS}V(n+1,0) - T^{(u)}_{AS}V(n,1) \ =\ c(n+1)\ +\ \frac{1}{\tilde{\Lambda}+\theta}\,(\Lambda(n)-\mu)c_B \ =\ c_B\,,$$

for n in the same range. This means that $T^{(u)}_{AS}V$ satisfies P4. □

Lemma 6 (Propagation of Picx for $T_{AS}^{(u)}$). *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies properties Picx and P4. Then $T_{AS}^{(u)}V$ satisfies Picx.*

Proof. Since V satisfies P4, we can apply (21) and (22) of Lemma 3 for all $n \leq N$. Proving that $T_{AS}^{(u)}V$ satisfies Picx amounts to proving two properties on $n \mapsto \Delta_n T_{AS}^{(u)}V(n,0)$: it should be non negative and increasing in the required intervals.

i) Non negativity. All the terms in the right-hand side of (21) and (22) are positive, since $\Delta_n V(n,0)$ is positive when $0 \leq n \leq N$, by Picx.

ii) Monotonicity. In the case $N = 1$, the monotonicity requirement of $\Delta_n T_{AS}^{(u)}V(n,0)$ vanishes. We therefore assume $N \geq 2$ in the following.

The first line of the right-hand side of (21) is an increasing function of n for $1 \leq n \leq N$, since V satisfies Picx. The second and the third lines are also increasing for $1 \leq n \leq N$, since V satisfies Picx, and therefore P2 and P3 by Lemma 1.

There remains to prove the increasing of increments at $n = 0$, that is equivalent to show that $\Delta_n T_{AS}^{(u)}V(0,0) \leq \Delta_n T_{AS}^{(u)}V(1,0)$. We obtain $\Delta_n T_{AS}^{(u)}V(0,0)$ from (22) and $\Delta_n T_{AS}^{(u)}V(1,0)$ from (21) evaluated at $n = 1$, since we have assumed that V satisfies P4. We obtain therefore, with $\alpha(0) = 0$, $\alpha(1) = \alpha$ and $\lambda(0) = \Lambda$:

$$\begin{aligned}(\tilde{\Lambda}+\theta)(\Delta_n T_{AS}^{(u)}V(0,0) - \Delta_n T_{AS}^{(u)}V(1,0)) \;=\;& c_B(\mu+\theta) - c_Q - \mu\Delta_n V(0,0)\\ &+\Lambda\Delta_n V(1,0) - \lambda(1)\Delta_n V(2,0)\\ &+N\alpha\Delta_n V(0,0) - (N\alpha-\alpha)\Delta_n V(1,0)\,.\end{aligned}$$

Adding and subtracting the quantities $\lambda(1)\Delta_n V(1,0)$, $\alpha\Delta_n V(0,0)$ and $\theta\Delta_n V(0,0)$ we have:

$$\begin{aligned}&(\tilde{\Lambda}+\theta)(\Delta_n T_{AS}^{(u)}V(0,0) - \Delta_n T_{AS}^{(u)}V(1,0))\\ &= c_B(\mu+\theta) - c_Q - \mu\Delta_n V(0,0) + \theta\Delta_n V(0,0) - \theta\Delta_n V(0,0)\\ &\quad +\Lambda\Delta_n V(1,0) - \lambda(1)\Delta_n V(2,0) - \lambda(1)\Delta_n V(1,0) + \lambda(1)\Delta_n V(1,0)\\ &\quad +N\alpha\Delta_n V(0,0) - (N\alpha-\alpha)\Delta_n V(1,0) + \alpha\Delta_n V(0,0) - \alpha\Delta_n V(0,0)\\ &= (\mu+\theta)(c_B - \Delta_n V(0,0)) - c_Q + (\Lambda-\lambda(1))\Delta_n V(1,0) + (\alpha+\theta)\Delta_n V(0,0)\\ &\quad +\lambda(1)(\Delta_n V(1,0) - \Delta_n V(2,0))\\ &\quad +(N\alpha-\alpha)(\Delta_n V(0,0) - \Delta_n V(1,0))\,.\end{aligned}$$

The two last terms in this expression are negative since $\Delta_n V(n,0)$ is increasing for $0 \leq n \leq N$ (the assumption that $N \geq 2$ is needed here). By assumption P1 b), $c_B - \Delta_n V(0,0) \leq 0$, then $(\mu+\theta)(c_B - \Delta_n V(0,0))$ is negative. For the remaining terms, observe that $\Lambda - \lambda(1) = \Lambda/N$ and that, by assumption P1 a), $\Delta_n V(n,0) \leq c_Q/(\alpha+\theta+\Lambda/N)$ for $n = 0, 1$. Therefore,

$$-c_Q + (\Lambda-\lambda(1))\Delta_n V(1,0) + (\alpha+\theta)\Delta_n V(0,0) \leq -c_Q + \frac{c_Q}{\alpha+\theta+\Lambda/N}(\alpha+\theta+\Lambda/N) = 0.$$

Summing up, we have shown that $\Delta_n T_{AS}^{(u)}V(0,0) - \Delta_n T_{AS}^{(u)}V(1,0) \leq 0$, so that $\Delta_n T_{AS}^{(u)}V(n,0)$ is increasing for $0 \leq n \leq N$.

We have therefore proved the non-negativity and the increasing of increments of $T_{AS}^{(u)}V$, hence the statement. □

Lemma 7 (Propagation of P1 b) for $T_{AS}^{(u)}$). *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies property Picx, P1 b) and P4. Then $T_{AS}^{(u)}V$ also satisfies P1 b).*

Proof. We proceed with bounding below the terms in (22), which holds since P4 is assumed. By Assumption P1 b) on V, $V(0,0) \geq c_B$. By Assumption Picx (the convexity part), $\Delta_n V(1,0) \geq \Delta_n V(0,0) \geq c_B$. Therefore,

$$(\tilde{\Lambda} + \theta)\Delta_n T_{AS}^{(u)} V(0,0) \ \geq \ c_B(\mu + \theta + \Lambda + \Phi) \ = \ c_B(\tilde{\Lambda} + \theta) \ .$$

We have proved that $T_{AS}^{(u)} V$ satisfies P1 b). □

Lemma 8 (Propagation of P1 a) for $T_{AS}^{(u)}$). *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies properties P1 a) and P4. Then $T_{AS}^{(u)} V$ also satisfies P1 a).*

Proof. Expression (21) holds for $T_{AS}^{(u)}$ for $1 \leq n \leq N$, since we have assumed P4. We proceed with bounding above the terms in (21). As a preliminary, observe that the bound

$$\lambda(n)\Delta_n V(n+1,0) \ \leq \ \frac{c_Q}{\alpha + \theta + \Lambda/N}\, \lambda(n)$$

holds for all $n \geq 0$. Indeed, it holds for $0 \leq n < N$ by P1 a) on V, but also for $n \geq N$ since, by Assumption 1, $\lambda(n) = 0$. In that case, both sides are 0. Then, all multiplying coefficients of $\Delta_n V$ in (21) being positive, we get, for $1 \leq n \leq N$,

$$\begin{aligned}
(\tilde{\Lambda} + \theta)\Delta_n T_{AS}^{(u)} V(n,0) &\leq c_Q + \frac{c_Q}{\alpha + \theta + \Lambda/N}\Big(\mu + \lambda(n) + \Lambda - \lambda(n-1) + (N-n)\alpha + (n-1)\alpha\Big) \\
&= \frac{c_Q}{\alpha + \theta + \Lambda/N}\ \big(\alpha + \theta + \Lambda/N + \mu - \Lambda/N + \Lambda + N\alpha - \alpha\big) \\
&= \frac{c_Q}{\alpha + \theta + \Lambda/N}\ (\tilde{\Lambda} + \theta)\ .
\end{aligned}$$

In other words, $T_{AS}^{(u)} V$ satisfies P1 a) restricted to $1 \leq n \leq N$.

There remains to handle the case $n = 0$. We then turn to (22). Bounding above, we have

$$\begin{aligned}
(\Lambda + \Phi + \mu + \theta)\Delta_n T_{AS}^{(u)} V(0,0) &= c_B\ (\mu + \theta) + \Lambda\Delta_n V(1,0) + \Phi\Delta_n V(0,0) \\
&\leq c_B\ (\mu + \theta) + (\Lambda + \Phi)\frac{c_Q}{\alpha + \theta + \Lambda/N} \\
&= \frac{c_B\ (\alpha + \theta + \Lambda/N)\ (\mu + \theta) + (\Lambda + \Phi)c_Q}{\alpha + \theta + \Lambda/N} \\
&\leq \frac{c_Q\ (\mu + \theta) + (\Lambda + \Phi)c_Q}{\alpha + \theta + \Lambda/N}\ .
\end{aligned} \tag{29}$$

In the last inequality, we have used (28) which is equivalent to $c_B(\alpha + \theta + \Lambda/N) \leq c_Q$. Inequality (29) implies $\Delta_n T_{AS}^{(u)} V(0,0) \leq c_Q/(\alpha + \theta + \Lambda/N)$, hence the lemma. □

We conclude the section with the proof of its main result:

Proof of Proposition 3. Since V satisfies Picx and P1 a), P1 b) and P4, $T_{AS}^{(u)} V$ satisfies P4 by Lemma 5, Picx by Lemma 6, P1 b) by Lemma 7 and P1 a) by Lemma 8. □

5.5 Invariant property for operator $T_{NS}^{(u)}$

Throughout this section, we shall assume that

$$c_B \;\geq\; \frac{c_Q}{\alpha+\theta+\Lambda/N}\;. \tag{30}$$

The next lemma states essentially that, under this condition, the fixed point of operator $T_{NS}^{(u)}$ has a linear part for $n \leq N$.

Lemma 9 (Fixed point property for $T_{NS}^{(u)}$). *Let Assumption 1 hold, and let $V \in \mathcal{V}$ be a function that satisfies Plin. Then $T_{NS}^{(u)}V$ also satisfies Plin. In other terms, $T_{NS}^{(u)}V(n,b) = V(n,b)$ for all $0 \leq n \leq N$, $b \in \{0,1\}$.*

Proof. Let us use in this proof the form $V(n,0) = A(n+B)$, valid for $0 \leq n \leq N$. Observe that since $\lambda(n)$ vanishes for $n \geq N$ by Assumption 1, we have $\lambda(n)V(n+1,0) = \lambda(n)A(n+1+B)$ for every $n \in \mathbb{N}$ (a "trick" that was used already in the proof of Lemma 8). Also, $\alpha(n)V(n-1,0) = \alpha(n)A(n-1+B)$ for all $1 \leq n \leq N+1$. From Equation (11) we then get

$$\begin{aligned}(\tilde{\Lambda}+\theta)T_{NS}^{(u)}V(n,0) &= nc_Q + \tilde{\Lambda}A(n+B) + \lambda(n)A - \alpha(n)A \\ &= nc_Q + \tilde{\Lambda}A(n+B) + \theta A(n+B) - \theta A(n+B) + \lambda(n)A - \alpha(n)A \\ &= (\tilde{\Lambda}+\theta)V(n,0) + nc_Q + A[-\theta(n+B) + \lambda(n) - \alpha(n)]\,.\end{aligned}$$

Let us replace A, B, $\lambda(n)$ and $\alpha(n)$ by their values. For $0 \leq n \leq N$, we get

$$\begin{aligned}(\tilde{\Lambda}+\theta)T_{NS}^{(u)}V(n,0) &= (\tilde{\Lambda}+\theta)V(n,0) + nc_Q + A\left(-\theta n - \Lambda + \Lambda - n\frac{\Lambda}{N} - n\alpha\right) \\ &= (\tilde{\Lambda}+\theta)V(n,0) + nc_Q - \frac{nc_Q}{\alpha+\theta+\Lambda/N}\left(\theta + \frac{\Lambda}{N} + \alpha\right) = (\tilde{\Lambda}+\theta)V(n,0)\end{aligned}$$

hence the statement for $b = 0$.

When $b = 1$, we see from Equation (12) that if $V(n,0) = V(n,1)$ for $0 \leq n \leq N$, then the terms multiplied by μ disappear and (12) reduces to (11). The previous calculation applies and $T_{NS}^{(u)}V(n,1) = V(n,0)$, hence the result. □

6 Optimality results

In this section, we collect optimality results for the approximated models and then we apply them to find the optimal policy for the Base Model (Definition 2). To this aim, we use the recent results of [8] allows us to extend the results in truncated model in the Base Model.

6.1 Optimality results in the approximated models

In this section, we study and exhibit the optimal policy in approximated models defined by Assumption 1. This is done using the propagation framework presented by Puterman [22] or Koole [17]. This framework does not work in the Base Model (as it will be seen in Section 7.1) but when adapted in truncated and smoothed models this framework allows to determine properties of optimal policies.

However, the propagation framework proofs usually proceed by showing the submodularity of the operator $T^{(u)}$ since the submodularity induces the monotonicity of the function $\Delta_q T^{(u)} V(n)$ and therefore the optimal decision rule. But, in our case it will be shown in Example 4, that the submodularity does not occur and we should refine the study of the sign of $\Delta_q T^{(u)} V(n)$.

Before proceedings, we comment briefly on the validity of using a framework from [22] in the context of the control model of [13] or [9]. Indeed, in the former control policies are applied at certain *decision epochs* and in the latter, they are applied continuously. However, since for finite Markov models it is known that optimal policies (in the sense of [22]) can be found in the set of Markov Deterministic Policies (denoted with D^{MD} in [22]), it does not matter whether the policy is seen as applied at some time instant or continuously: as long as the state of the system does not change, the policy does not change. Moreover, structure propagation is essentially concerned with value functions and Bellman operators. As long it has been proved, with either MDP model, that the optimal value function satisfies some Bellman equation, the framework may be used, in line with the roadmap proposed in [9].

6.1.1 Overview of the propagation framework

We give now a brief recall of the propagation framework. Consider a MDP with discrete state space and *bounded* rewards, with the expected total discounted cost as optimization criterion. Denote with $\mathcal{V}$ the set of value functions. Let $d \in F$ be a Markov (feedback) decision rule, let T_d be the operator acting on $\mathcal{V}$ that computes the one-step value of the policy d with terminal value $V \in \mathcal{V}$. Let $T = \inf_{d \in F} T_d$ be the dynamic programming operator of the problem.

In the following statement, $\mathcal{V}_\sigma \in \mathcal{V}$ is interpreted a set of value functions such that some "structured" properties are satisfied, and $\mathcal{D}_\sigma$ is a corresponding set of "structured" decision rules.

Theorem 3 (Theorem 6.11.1 in [22])**.** *Consider a countable-state, discrete-time Markov Decision Problem, with bounded costs, and dynamic programming operator T. Assume that:*

1. *for each $V \in \mathcal{V}$, there exists a deterministic Markov decision rule d such that $TV = T_d V$,*

and that there exist non-empty sets $\mathcal{V}_\sigma \subset \mathcal{V}$ and $\mathcal{D}_\sigma \subset F$ such that:

2. *$V \in \mathcal{V}_\sigma$ implies $TV \in \mathcal{V}_\sigma$,*
3. *$V \in \mathcal{V}_\sigma$ implies there exists a decision rule d' such that $d' \in \mathcal{D}_\sigma \cap \arg\min_d T_d V$,*
4. *$\mathcal{V}_\sigma$ is a closed subset of $\mathcal{V}$ under simple convergence.*

Then, there exists an optimal stationary policy d^* in F with $d^* \in \mathcal{D}_\sigma \cap \arg\min_d T_d V$.

We will apply this theorem to the approximated models specified by Assumption 1. In this context, the set $\mathcal{V}_\sigma$ will typically be defined by some of the properties introduced in Definition 7, and the set $\mathcal{D}_\sigma$ will specify which action to take in certain states of $\mathcal{X}$.

We identify the suitable sets and prove that they satisfy the assumptions of Theorem 3 in Section 6.1.3. A technical point is to check that the sets $\mathcal{V}_\sigma$ and $\mathcal{D}_\sigma$ are not empty. We address this point in 6.1.4. Before this, we establish in Section 6.1.2 a preliminary result that relates the dynamic programming operator $T^{(u)}$ to the operators $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$.

6.1.2 Study of the sign of $\Delta_q T^{(u)}$

Now, we establish formally results about the sign of the difference $\Delta_q T^{(u)} V(n) = T^{(u)}_{AS} V(n,0) - (T^{(u)}_{NS} V)(n,0)$ which imply partial optimality results for the approximated models. More precisely, the two following lemmas allows to handle requirement 3 of Theorem 3.

Lemma 10 (Result for small c_B). *Let Assumption 1 and (28) hold. Let $V \in \mathcal{V}$ be a function with properties Picx, P1 a), P1 b) and P4. Then, for $0 \le n \le N+1$:*

i) $\Delta_q T^{(u)} V(n) \le 0$*;*

ii) $T^{(u)} V(n,0) = T^{(u)}_{AS} V(n,0)$.

Proof. *i*) Since V satisfies P4, Lemma 4 applies. Then $\Delta_q T^{(u)} V(0) = 0$. Further, under Assumption (28) we have Equation (24), namely:

$$(\tilde{\Lambda} + \theta)\Delta_q T^{(u)} V(n) = (\mu + \theta)\,(c_B - \Delta_n V(n-1,0)) - c_Q \\ + (\alpha + \theta)\ \Delta_n V(n-1,0) + (\Lambda/N)\ \Delta_n V(n,0)\ ,$$

for $1 \le n \le N$. By Assumption P1 b), we have $c_B - \Delta_n V(n-1,0) \le 0$. By Picx, we have $\Delta_n V(n-1,0) \le \Delta_n V(n,0)$. By P1 a) it follows that $\Delta_n V(n-1,0)$ and $\Delta_n V(n,0)$ are both smaller than $c_Q/(\alpha + \theta + \Lambda/N)$. Used in (24), these bounds lead to $(\tilde{\Lambda} + \theta)\Delta_q T^{(u)} V(n) \le 0$ for $1 \le n \le N$.

Finally, for $n = N+1$ we have Equation (25) from which it follows:

$$(\tilde{\Lambda} + \theta)\Delta_q T^{(u)} V(N+1) = c_B\,(\mu + \theta) - c_Q - \mu \Delta_n V(N,0)\,.$$

By (28), we have $c_B\,(\mu + \theta + \Lambda/N) - c_Q \le 0$, hence $c_B(\mu + \theta) - c_Q \le -c_B \Lambda/N \le 0$. Since $\Delta_n V(N,0) \ge 0$ by Picx, then $\Delta_q T^{(u)} V(N+1) \le 0$. This completes the proof of *i)*.

ii) From *i*), $\Delta_q T^{(u)} V(n) \le 0$ for $0 \le n \le N+1$, in other words, $T^{(u)}_{AS} V(n,0) \le (T^{(u)}_{NS} V)(n,0)$. But according to Definition 6, $T^{(u)} V = \min\{T^{(u)}_{AS} V, T^{(u)}_{NS} V\}$. We conclude that for $0 \le n \le N+1$, $T^{(u)} V(n,0) = T^{(u)}_{AS} V(n,0)$. □

We have the similar result for large c_B.

Lemma 11 (Result for large c_B). *Let Assumption 1 and Condition* (30) *hold. Let V be a function satisfying Plin. Then, for $0 \le n \le N$:*

i) $\Delta_q T^{(u)} V(n,0) \ge 0$ *;*

ii) $T^{(u)} V(n,0) = T_{NS} V(n,0)$.

Proof. i) According to Lemma 4 i) and iv): for a function V that satisfies Plin then on the one hand, $\Delta_q T^{(u)} V(0) = 0$ and on the other hand, (26) holds, namely, for all $1 \le n \le N$:

$$(\tilde{\Lambda} + \theta)\Delta_q T^{(u)} V(n,0) = \big(\mu + \theta + \lambda(n-1) + \alpha(n-1)\big)\left(c_B - \frac{c_Q}{\alpha + \theta + \Lambda/N}\right)\ .$$

Under Condition (30), this is the product of two non-negative terms, which proves i).

ii) From *i*), $\Delta_q T^{(u)} V(n) \ge 0$ for $0 \le n \le N$, in other words, $T^{(u)}_{AS} V(n,0) \ge (T^{(u)}_{NS} V)(n,0)$. Since, according to Definition 6, $T^{(u)} V = \min\{T^{(u)}_{AS} V, T^{(u)}_{NS} V\}$, we conclude that for $0 \le n \le N$, $T^{(u)} V(n,0) = T^{(u)}_{NS} V(n,0)$. This concludes the proof. □

6.1.3 Invariant properties for operator $T^{(u)}$

We now integrate the results for operators $T^{(u)}_{AS}$ and $T^{(u)}_{NS}$ to obtain propagation results for operator $T^{(u)}$. The objective is to ascertain requirement 2 of Theorem 3.

Proposition 4 (Propagation for $T^{(u)}$ for small c_B)**.** *Let Assumption 1 and Condition (28) hold. Let $V \in \mathcal{V}$ be a function which satisfies properties Picx, P1 a), P1 b) and P4. Then $T^{(u)}V$ also satisfies these four properties.*

Proof. According to Lemma 10, $T^{(u)}V(n,0) = T^{(u)}_{AS}V(n,0)$ for all $0 \leq n \leq N+1$. On the other hand, we have, by definition, $T^{(u)}_{AS}V(n,1) = T^{(u)}_{NS}V(n,1)$ for all $n \geq 0$. Therefore $T^{(u)}V(n,1) = T^{(u)}_{AS}V(n,1)$ for all $n \geq 0$, in particular for $0 \leq n \leq N+1$.

By Proposition 3, $T^{(u)}_{AS}V$ has the four properties Picx, P1 a), P1 b) and P4. All properties involve values of the "V" function at (n,b) with $0 \leq n \leq N+1$. If they hold for $T^{(u)}_{AS}V$, they hold for $T^{(u)}V$. □

Proposition 5 (Propagation for $T^{(u)}$ for large c_B)**.** *Let Assumption 1 hold, Condition* (30) *hold and let $V \in \mathcal{V}$ be a function which satisfies Condition* (17)*. Then $T^{(u)}V$ also satisfies Condition* (17)*.*

Proof. According to Lemma 11, $T^{(u)}V(n,0) = T^{(u)}_{NS}V(n,0)$ for all $0 \leq n \leq N$. On the other hand, we have, by definition, $T^{(u)}_{NS}V(n,1) = T^{(u)}_{NS}V(n,1)$ for all $n \geq 0$. Therefore $T^{(u)}V(n,1) = T^{(u)}_{NS}V(n,1)$ for all $n \geq 0$, in particular for $0 \leq n \leq N$.

By Lemma 9, we know that a function V that satisfies (17) is such that $T^{(u)}_{NS}V$ also satisfies it. Since $T^{(u)}_{NS}V$ and $T^{(u)}V$ coincide for states $(n,0)$, $0 \leq n \leq N$, $T^{(u)}$ satisfies (17). □

6.1.4 Initial Value function

The structural propagation framework is based on the propagation of properties through the dynamic programming operator. It works if it is possible to find a function in order to initiate the process. This necessity is present in the requirement that the set of functions $\mathcal{V}_\sigma$ be non-empty in Theorem 3.

In this section, we identify a function that indeed satisfies the properties to be propagated, namely, Picx, P1 a), P1 b), and P4 when (28) holds or Equation (17) when (30) holds.

Observe that the very usual choice of a null function does not work here: Condition P1 b) coupled with Picx imposes that the function be strictly increasing and convex. Up to our knowledge it is an unusual case. We exhibit now such a function which satisfies all the required properties. We chose a form that covers both cases (28) and (30).

Let us define

$$\chi = (\alpha + \theta + \frac{\Lambda}{N})\frac{c_B}{c_Q} - 1 \,,$$

and let us define $V_0(n,b)$ for all $n \geq 0$:[6]

$$V_0(n,0) = \frac{c_Q}{\alpha+\theta+\Lambda/N}\left(n + \frac{\max(0,\chi)}{\chi}\frac{\Lambda}{\theta}\right) \tag{31}$$

$$V_0(n,1) = \frac{c_Q}{\alpha+\theta+\Lambda/N}\left(n + \frac{\max(0,\chi)}{\chi}\frac{\Lambda}{\theta} + \min(\chi,0)\right). \tag{32}$$

We now state that the above function satisfies all properties to be propagated, that is:

[6] In the case $\chi = 0$, the term $\max(0,\chi)/\chi$ is to be interpreted as 0.

Lemma 12. *Let Assumption 1 hold. Then:*

i) If (28) *holds, the function* V_0 *satisfies properties Picx, P1 a), P1 b) and P4.*

ii) If (30) *holds, the function* V_0 *satisfies Plin.*

Proof. $i)$ When (28) holds (small c_B) then $\chi \le 0$ and we have, for all $0 \le n \le N$,

$$\Delta_n V_0(n,0) = V_0(n+1,0) - V_0(n,0) = \frac{c_Q}{\alpha + \theta + \Lambda/N}.$$

Then Picx and P1 a) are clearly satisfied, and since $c_B \le \frac{c_Q}{\alpha+\theta+\Lambda/N}$ then P1 b) is satisfied as well. We also have

$$V_0(n+1,0) - V_0(n,1) = \frac{c_Q}{\alpha + \theta + \Lambda/N}(1+\chi) = c_B$$

which is P4.

$ii)$ When (30) holds (large c_B) then $\chi \ge 0$ and the definition of $V_0(n,b)$ corresponds with (17) for all $0 \le n \le N$ and $b \in \{0,1\}$. □

6.1.5 Optimal policy for approximate models

We prove now that the optimal policy in approximate models is a threshold policy with a threshold equal to one.

Theorem 4. *Let Assumption 1 hold.*

i) If (28) holds, then there exists an optimal feedback policy γ_N^* *such that* $\gamma_N^*(n,0) = 1$ *for all* $1 \le n \le N+1$.

ii) If (30) holds, then there exists an optimal feedback policy γ_N^* *such that* $\gamma_N^*(n,0) = 0$ *for all* $1 \le n \le N$.

Proof. $i)$ We apply Theorem 3 with the sets $\mathcal{V}_\sigma$ and $\mathcal{D}_\sigma$ defined as:

$$\mathcal{V}_\sigma = \{V \in \mathcal{V} \mid V \text{ satisfies Picx, P1 a), P1 b) and P4}\}$$

and

$$\mathcal{D}_\sigma = \{\pi \in F \mid \pi(n,0) = 1, 0 \le n \le N+1\}.$$

The set $\mathcal{V}_\sigma$ is not empty since it contains V_0, defined in Section 6.1.4, by virtue of Lemma 12. The assumption 1) of Theorem 3 holds because all $\mathcal{A}_x$ are finite ([22, Theorem 6.2.10]). That $V \in \mathcal{V}_\sigma$ implies $T^{(u)}V \in \mathcal{V}_\sigma$ is guaranteed by Proposition 4. That $V \in \mathcal{V}_\sigma$ implies the existence of some $\pi \in D^u$ with the desired property, is a consequence of Lemma 10. Finally, the closure of $\mathcal{V}_\sigma$ under simple convergence is clear, since this set is defined by a finite number of linear constraints.

$ii)$ We now apply Theorem 3 with the sets $\mathcal{V}_\sigma$ and $\mathcal{D}_\sigma$ defined as:

$$\mathcal{V}_\sigma = \{V \in \mathcal{V} \mid V \text{ satisfies Plin}\}$$

and

$$\mathcal{D}_\sigma = \{\pi \in F \mid \pi(n,0) = 0, 0 \le n \le N\}.$$

The set $\mathcal{V}_\sigma$ is clearly not empty. The assumption 1) of Theorem 3 holds as above. That $V \in \mathcal{V}_\sigma$ implies $T^{(u)}V \in \mathcal{V}_\sigma$ is guaranteed by Proposition 5. That $V \in \mathcal{V}_\sigma$ implies the existence of some $\pi \in \mathcal{D}^u$ with the desired property, is a consequence of Lemma 11. Finally, the closure of $\mathcal{V}_\sigma$ under simple convergence is clear. □

6.2 Optimality in the base model

In this section, we proceed with the proof of Theorem 1. We shall apply Theorem 5.1 of [8], included as Theorem 8 in Appendix C for the sake of self-containedness. For this, we need to check all parts of Assumption 5.

Proof of Theorem 1. In the context of these results, recall the notation $A = \mathcal{N} \times F$ for the parameter set: the models $X(a) = X(N, \pi)$ are parametrized by the truncation parameter N and the stationary deterministic policy π. Select a real number $\gamma \in (0, \theta)$, complying with Assumption $5h$).[7]

For every $a = (N, \pi) \in A$, the induced Markov process $X(a)$ has bounded transition rates so that the transition matrices are conservative. That the Markov chains are stable for all δ derives from the fact that the chain obtained with the "No Service" policy is always stable. Indeed: this chain is a birth-death process with population-dependent birth rates which vanish for populations larger than N. A stochastic comparison argument is then used for other policies. Assumption $5a$) therefore holds.

The set A is discrete, because F itself is discrete, since action spaces are finite. Hence A is locally compact: Assumption $5b$) holds.

According to their definitions in Section 2.2, the transition rates are either constant, or simple combinations of $\lambda(n)$ and $\alpha(n)$, themselves continuous functions of n when defined in Assumption 1. Assumption $5c$) therefore holds. Similarly with cost functions defined in Definition 1 where the function $k(\cdot)$ is defined (3): Assumption $5f$) also holds.

According to Lemma 15, for every γ the function $V = w_\varepsilon$ with $w_\varepsilon(n, b) = e^{n\varepsilon}$ and $\varepsilon \leq \log(1 + \gamma/\Lambda)$, is such that Assumption $5d$) holds. Select then $V = w_\varepsilon$ and $W = w_{\varepsilon'}$ with $0 < \varepsilon < \varepsilon' \leq \log(1 + \gamma/\Lambda)$. Then $W(n, b)/V(n, b)$ is unbounded and this implies the existence of a sequence of sets K_n as in Assumption $5e$), choosing the value $\theta = \gamma$.

Still according to the definition of the approximated models, we have, for $x = (n, b)$, $\sup_a c_x(a)$ is a polynomial in n, and consequently, $\sup_{x,a} c_x(a)/V(x) < \infty$. This checks Assumption $5g$).

Finally, the "product property" of Assumption $5i$) results from the fact that the approximation of Assumption 1 is "independent" from the stationary deterministic policy d. Indeed, consider a fixed $N \in \mathcal{N}$ and the family of models $\{X(N, \pi); \pi \in F\}$. Then $\mathcal{F}$ can be represented as $F = \prod_{s \in \mathcal{X}} A_x$ with $A_x = \mathcal{N} \times \mathcal{A}_x$, where $\mathcal{A}_x$ is defined in Section 2.2. The sets A_x are finite hence compact. Moreover, for some $\pi \in F$, consider the rate matrix $Q(N, \pi)$ obtained by applying the approximation factor N and the control policy π. The product property holds if the row x of this matrix depends only on $\pi(x)$. But this row is given by the rates defined in Section 2.2 with the functions $\alpha(.)$ and $\lambda(.)$ specified in Assumption 1. These values clearly depend only on $\pi(x)$.

All requirements of Assumption 5 are therefore satisfied.

Consider first the case where $c_Q > c_B(\alpha + \theta)$. Then there exists a $N_0 \in \mathbb{N}$, $N_0 \geq 1$, such that, for all $N \geq N_0$, Condition (28) holds.[8] According to Theorem 4 i), for every such N there is an optimal policy π^*_N where $\pi^*_N(n, 0) = 1$ for $0 \leq n \leq N$. Whatever the actual values of $\pi^*_N(n, 0)$ for $n > N \geq N_0$, or those of policies π^*_N for $N < N_0$, this sequence of policies has the accumulation point π_{AS} where $\pi_{AS}(n, 0) = 1$ for all $n \in \mathbb{N}$, is the AS policy introduced in Definition 3. In other words, AS is optimal and Theorem 1 b) is proved.

Assume now that $c_Q \leq c_B(\alpha + \theta)$. For all $N \in \mathbb{N}$, Condition (30) holds. Starting with Theorem 4 ii), the previous reasoning allows to conclude that π_{NS} is optimal. Theorem 1 a) is proved, as well as the statement in Theorem 1 c) that π_{NS} is optimal when $c_Q = c_B(\alpha + \theta)$.

[7] Observe that in the statement Assumption 5, the discount factor is called α.

[8] Namely: $N_0 = \lceil c_B \Lambda / (c_Q - c_B(\alpha + \theta)) \rceil$.

There remains to be proved that π_{AS} is also optimal when $c_Q = c_B(\alpha + \theta)$. To that end, consider the family of MDPs of the Base Model, parametrized by c_Q. More precisely, and in order to match the framework of Theorem 8, consider the family of models parametrized by N with $c_Q^{(N)} = c_B(\alpha + \theta)(1 - 1/N)$, all other parameters being unchanged. The conditions of Assumption 5 are satisfied: either they were already checked with the Base Model for features involving only transition rates, or they are straightforward to check for features involving costs. It is then possible to apply again Theorem 8 to conclude that since π_{AS} is optimal for all $N \in \mathbb{N} \setminus \{0\}$, it is also optimal for $N = \infty$, that is, when $c_Q = c_B(\alpha + \theta)$. □

We conclude this section with a complement on optimal value functions, also an application of Theorem 8.

Theorem 5. *The value function V^* of the optimization problem in the Base Model:*

i) is increasing, convex, is such that $\Delta_n V^*(0, 0) \geq c_B$ and $\Delta_n V^*(n, 0) \leq c_Q/(\alpha + \theta)$ for all $n \geq 0$;

ii) when $c_Q \leq c_B(\alpha + \theta)$, is given by, for all $(n, b) \in \mathcal{X}$:

$$V^*(n, b) = V_{NS}(n, b) = \frac{c_Q}{\alpha + \theta}\left(n + \frac{\Lambda}{\theta}\right) . \tag{33}$$

Proof. When $c_Q \leq c_B(\alpha + \theta)$, Theorem 1 says that γ_{NS} is optimal, so that $V^* = V_{NS}$. That the function defined in (33) is indeed V_{NS} is checked by checking that it is a fixed point of the "no serve" equation, that is (see Proposition 1 and Equation (6)): for all $n \geq 0$,

$$V(n, 0) = \frac{1}{\lambda(n) + \alpha(n) + \theta}\left[nc_Q + \lambda(n)V(n + 1, 0) + \alpha(n)V(n - 1, 0)\right] .$$

In addition, under γ_{NS}, $V^*(n, 0) = V^*(n, 1)$. This proves claim $ii)$.

The function V_{NS} has all properties listed in $i)$ when $c_Q \leq c_B(\alpha + \theta)$ since $\Delta_n V_{NS}(n, 0) = c_Q/(\alpha + \theta) \geq c_B$. When $c_Q > c_B(\alpha + \theta)$, we use the approximated models. Theorem 8 $i)$ applies and V^* is the limit of a sequence of approximated value functions V_N. Each of these has properties Picx, P1 a), P1 b) and P4 according to the proof of Theorem 4. In the limit $N \to \infty$, V^* has therefore the properties stated in claim $i)$.[9] □

7 Methodological comments

In this section we group a set of observations on the features of the problem solved in the paper. Each observation below details the technical points which make a traditional proof method ineffective.

7.1 Intractability of structural properties propagation framework in the Base Model

The classic approach in order to show that the optimal policies are threshold policies uses the framework of propagation of structural properties by the dynamic programming operator, recalled in Section 6.1.1. It needs to find a set of "structured" value functions, V^σ, which is invariant under this operator. The usual sets considered involve monotony an/or some form of convexity.

[9] This limiting argument may be used also in the case $c_Q \leq c_B(\alpha + \theta)$ with property Plin.

Example 1 exhibits cases in which simple properties like increasingness, convexity or concavity are not conserved during the value iteration procedure, when it uses the "natural" dynamic programming operator defined on the Base Model of Definition 2. This suggests that the structured policy propagation framework is likely to fail for this operator. As shown in Section 5, the operator resulting from smoothing and truncation does propagate convexity.

Example 1 (Non convexity of iterations of Value Iteration)**.** *Consider the Base Model with the following parameters:* $\Lambda = 0.5$*,* $\mu = 5$*,* $\alpha = 1$ *and* $\theta = 0.1$*; the costs are* $c_B = 1.0$*,* $c_L = 2.0$ *and* $c_H = 2.0$*. Computations are done in a finite model of size* $S = 100$*.*

We define V_0 *the initial value function. We compute* $V_{\ell+1}$ *with respect to* V_ℓ *using the Bellman operator defined from equations* (5)–(8) *i.e.:*

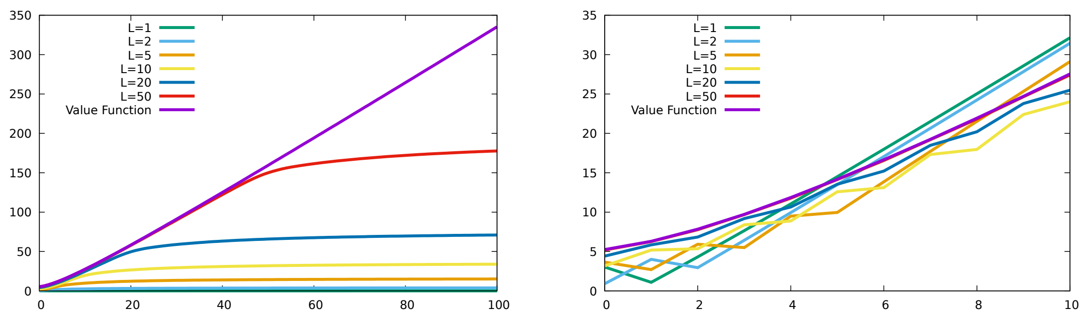
$$V_{\ell+1}(0,0) = \frac{1}{\Lambda+\theta}\left[k(0) + \Lambda V_\ell(1,0)\right],$$
$$V_{\ell+1}(n,0) = \min\Big\{c_B + \frac{1}{\Lambda+(n-1)\alpha+\mu+\theta}\big[k(n-1) + \Lambda V_\ell(n,1) + (n-1)\alpha V_\ell(n-2,1) + \mu V_\ell(n-1,0)\big], \frac{1}{\Lambda+n\alpha+\theta}\left[k(n) + \Lambda V_\ell(n+1,0) + n\alpha V_\ell(n-1,0)\right]\Big\}$$

for $n \geq 1$ *and*

$$V_{\ell+1}(n,1) = \frac{1}{\Lambda+n\alpha+\mu+\theta}\left[k(n) + \Lambda V_\ell(n+1,1) + n\alpha V_\ell(n-1,1) + \mu V_\ell(n,0)\right],$$

for $n \geq 0$*.*

Figure 1 represents the functions V_ℓ *for different values of* ℓ *as well as their numerical limit, obtained with Value Iteration until no evolution is observed. This is done for two initial values. For the curves on the left, we choose* $V_0(n,0) = V_0(n,1) = 0$ *for all* $n \geq 0$*. It is observed that the initial function is linear, that the first iterate is concave, that the following iterates are neither concave nor convex, while the limit is convex (Theorem 5). For the curves on the right-hand side, we choose the value function used in the proof of the paper (defined by Equations* (31) *and* (32)*). The initial value is increasing convex, then the first iterate is convex but not increasing, that the following iterates are neither increasing nor convex while the limit is increasing convex.*



(a) initial value null (b) initial value of the proof (zoom)

Figure 1: Plot of the iterations of the value function

7.2 Intractability of sample Path approaches

Due to the simple form of the policy it could be thought that a simpler method such as sample path should be sufficient. Example 2 shows that proving the result of Theorem 1 with the use of sample path arguments is not direct. The counterexample is constructed in the non uniformized case but a similar example can be constructed for the smoothed and uniformized case in the same way.

Example 2 (Sample Path comparison fails)**.** *We assume a case in which at the start there is only one single customer in the queue and no service in progress:* $x_0 = (1,0)$*. We also assume that the impatience of the customer is such that its departure time is* D_0 *if it is not served with furthermore* $D_0 \leq A_1$ *where* A_1 *is the epoch of the next arrival.*

If we choose to serve the customer at time 0 *then it costs* $J_{AS} = c_B + J_{A_1}$ *where* J_{A_1} *is the cost incurred by the system from time* A_1*. Otherwise, if we do not serve the customer then it costs*

$$J_{NS} = c_H \int_0^{D_0} e^{-\theta t} dt + c_L e^{-\theta D_0} + J_{A_1}.$$

We have

$$J_{NS} - J_{A_1} = c_H \frac{1 - e^{-\theta D_0}}{\theta} + c_L e^{-\theta D_0} = (1 - e^{-\theta D_0}) \frac{c_H}{\theta} + e^{-\theta D_0}\, c_L\,.$$

Observe that this is a convex combination, and as D_0 *ranges from* 0 *to* ∞*, this value can be any number in the interval* $I := [c_L, c_H/\theta)$ *or* $I := (c_H/\theta, c_L]$*, depending on the relative positions of the extremities.*

Let us assume that we have parameter values such that AS *is optimal. We will now exhibit some trajectories for which it is not optimal to serve the initial customer although always serving is optimal. The condition for* AS *to be optimal is (Theorem 1):*

$$c_B \;\leq\; \overline{c}_B := \frac{c_H + \alpha c_L}{\alpha + \theta} \;=\; \frac{c_H}{\theta}\,\frac{\theta}{\alpha + \theta} + c_L\,\frac{\alpha}{\alpha + \theta}\,.$$

This is also a convex combination, so that $\overline{c}_B$ *lies in the interval* I*. If* $J_{NS} - J_{A_1} \leq J_{AS} - J_{A_1} = c_B$ *then it is optimal* not *to serve the customer.*

Assume in addition that $c_B > \min\{c_L, c_H/\theta\}$*. Then if* $c_B \leq \overline{c}_B \in I$*,* c_B *necessarily belongs to the interior of* I*. We have argued that that there exists a* D_0 *such that* $J_{NS} - J_{A_1}$ *is anywhere we want in* I*, in particular smaller than* c_B*. Picking such a* D_0*, together with* $A_1 > D_0$*, we have the desired trajectory.*

The argument also suggests that the sample path comparison does work when $c_B \leq \min\{c_L, c_H/\theta\}$ since not serving the customer will incur a cost belonging to the interval I.

7.3 Submodularity study

This section is dedicated to the study of the submodularity property which is widely used in the proofs of optimality for threshold policies. The study is done differently for the base model and for the smoothed model since the properties of the operators involved are different.

The submodularity referred to here is the one involving the state $x = (n, b)$ and the decision $q \in \{0, 1\}$. The operator T_q on $\mathcal{V}$ is defined as the action of decision q on a function V. The analysis of [22, Section 4.7] relates the super- or submodularity of $(q, x) \mapsto T_q V(x)$ to the optimality of monotone policies, which are threshold policies in our case. In particular, Lemma 4.7.6 from [22] implies that if $\Delta_q TV(n, b) = T_1 V(n, b) - T_0 V(n, b)$ (see also (23)) is monotonous in n, then a threshold policy is optimal.

7.3.1 Details on submodularity in the Base Model

In the Base Model case, Example 3 strongly suggests that the usual structural propagation of submodularity is not verified as soon as impatience is present.

Example 3 (Lack of submodularity, Base Model)**.** *We study the difference between the two terms of the* $\min$ *in the Bellman Equation. We hence define, for* $n \geq 1$, $T_q V(n,b)$ *by* $T_q V(n,b) = qA + (1-q)B$ *where* A *is given by Equation* (5) *:*

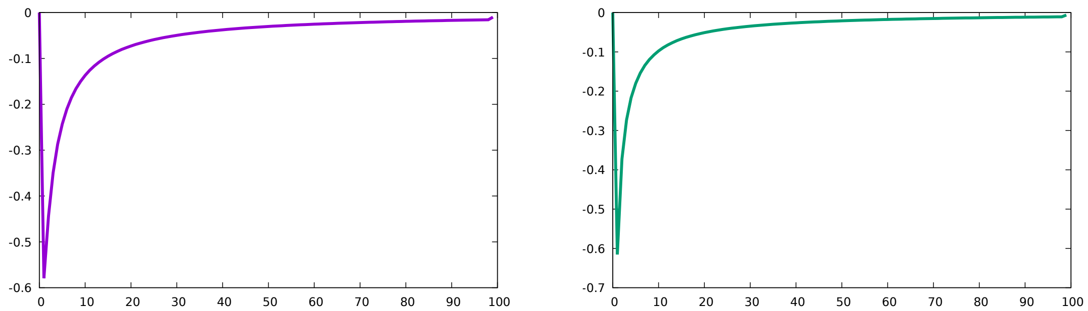
$$A = c_B + \frac{1}{\lambda(n-1) + \alpha(n-1) + \mu + \theta} \big[k(n-1) + \lambda(n-1)V(n,1) + \alpha(n-1)V(n-2,1) + \mu V(n-1,0) \big],$$

and B *is given by Equation* (6)*:*

$$B = \frac{1}{\lambda(n) + \alpha(n) + \theta} \left[k(n) + \lambda(n)V(n+1,0) + \alpha(n)V(n-1,0) \right] .$$

On Figure 2, we can observe the plotting of the difference $T_1 V(n,b) - T_0 V(n,b)$ *for two sets of parameters, where* V *is the function obtained by Value Iteration, see Example 1. It is observed that these ones are decreasing, then increasing from* $n \geq 1$*. This means that the function* $T_q V(n,b)$ *is neither sub- nor supermodular.*

Parameters have been chosen such that $\Lambda = 0.5$*, and* $\theta = 1.5$ *and such that the costs are* $c_B = 1.0$, $c_L = 2.0$ *and* $c_H = 2.0$*. For the left plot we have* $\mu = 2$ *and* $\alpha = 1.5$ *while for the right plot we have* $\mu = 1.5$ *and* $\alpha = 2$*. Computations are done in a finite model of size* 100.



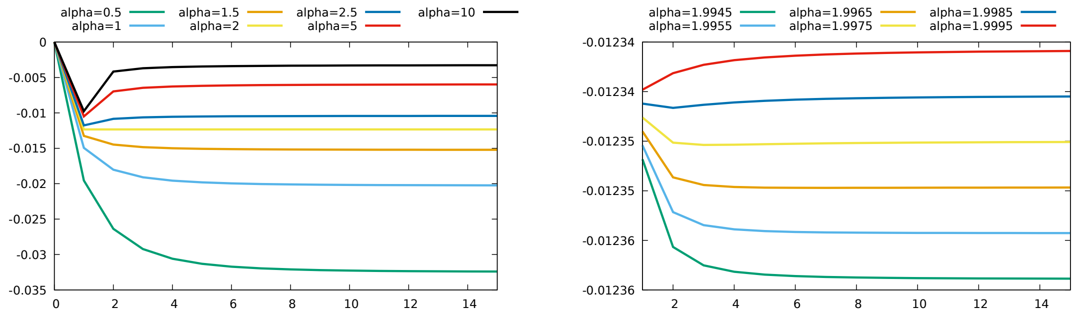
(a) $\mu - \alpha \geq 0$ (b) $\mu - \alpha \leq 0$

Figure 2: Non submodularity

7.3.2 Details on submodularity in approximated models

Example 3 indicates that there exists some cases for which the submodularity does not hold when we use the base model operator: we investigate this issue in the case of smoothed and truncated model. We show that first there exist cases where the submodularity does not hold with the base model operator while it holds for the operator of the approximated model. Then, we show for approximated models that, according to the relative values of the loss rate compared to the service rate, the dynamics of the queueing system is different. Hence, according to the dynamics, the operator can be either submodular or supermodular. This is illustrated on Figure 3.

Example 4. *On Figure 3, we plot the function $\Delta_q T^{(u)} V(n)$ (on the y-axis) with respect to n (on the x-axis) for different values of α. We choose the parameters such that Inequality* (28) *is satisfied for all the cases plotted. We fix $\Lambda = 0.5$, $\mu = 2.0$, $\theta = 1.5$ and $c_B = 1$, $c_H = 2$ and $c_L = 3$ and we let α vary between* 0.5 *and* 10*. Computation are done in a model where $N = 100$, implying $\Lambda/N = 0.005$.*



(a) (b)

Figure 3: Two dynamics when AS is optimal

In line with Inequality (28) *and Lemma 10, it can be seen that all the functions Δ_q are negative. Furthermore, it can be noticed that according to the sign of $\alpha - \mu$ the monotonicity is different. When $\alpha - \mu \geq 0$, the function $\Delta_q T^{(u)} V(n)$ is decreasing then increasing. When $\alpha - \mu < -\Lambda/N$, $\Delta_q T^{(u)} V(n)$ is decreasing, which is equivalent to submodularity of $T_q V$ as seen above. When $-\Lambda/N \leq \alpha - \mu < 0$, monotonicity may hold or does not hold, as can be seen on Figure 3 b).*

Example 4 relativizes the claims made, for example, in [12] and [17] that impatience breaks structural properties. Actually, only submodularity was broken, and there nevertheless exist some structural properties as seen in the previous parts of this work.

At last, it can be noticed that the left plot of Figure 2 corresponds with the third curve of Figure 3 since they have the same parameters $\mu = 2$ and $\alpha = 1.5$ in input. However, the two curves are obviously different. This illustrates again the structural difference of both dynamic programming operators, which are here applied to the same optimal value function which is their common fixed point.

7.3.3 Study of the change of structural properties

This change in the dynamics according to the sign of $\alpha - \mu$ is an intrinsic characteristic of the approximated model as we will show now.

Lemma 13. *Let Assumption 1 and* (28) *hold. Let V be the optimal value function of the smoothed and uniformized problem for some value of N. The function $\Delta_q T^{(u)} V(n)$ is:*

i) decreasing between $n = 0$ and $n = 1$;

ii) increasing for $1 \leq n \leq N$ when $\alpha - \mu \geq 0$.

Moreover,

iii) for every real number $\kappa > 0$, there exists some N_0 such that for every $N \geq N_0$, if a) $\alpha - \mu < -\Lambda/N$, and b) $\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0) \geq \kappa$ for $2 \leq n \leq N$, then the function $\Delta_q T^{(u)} V(n)$ is decreasing for $1 \leq n \leq N$.

Proof. Although this has not been stated explicitly yet, the fact that V satisfies Picx, P1 a), P1 b) and P4 is clear from the proof of Theorem 4, see also the proof of Theorem 5.
$i)$ Since $\Delta_q T^{(u)} V(n) = 0$, then $i)$ is a direct consequence of Lemma 10.
$ii)$ Since V satisfies P4, we can use Equation (27) from the proof of Lemma 4, namely, for $1 \leq n \leq N$:

$$(\tilde{\Lambda} + \theta)\Delta_q T^{(u)} V(n) = (\mu + \theta)\, c_B - c_Q + (\alpha - \mu)\ \Delta_n V(n-1, 0) + (\Lambda/N)\ \Delta_n V(n, 0)\,.$$

When $\alpha - \mu \geq 0$, then $\Delta_q T^{(u)} V(n)$ is the sum of a constant and two functions that are increasing for $1 \leq n \leq N$. The result is then increasing in the same range, which proves $ii)$.
$iii)$ When $\alpha - \mu < 0$ then $\Delta_q T^{(u)} V(n)$ is the sum of an increasing function and a decreasing function and the result may be non-monotone. However, the disturbance that results from the smoothing, i.e. the term with Λ/N as factor, may be made small in front of the decreasing term.

To that end, we use the above formula to get, for $2 \leq n \leq N$,

$$(\tilde{\Lambda} + \theta)\left(\Delta_q T^{(u)} V(n) - \Delta_q T^{(u)} V(n-1)\right) =$$
$$(\alpha - \mu)\big(\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0)\big) + \frac{\Lambda}{N}\big(\Delta_n V(n, 0) - \Delta_n V(n-1, 0)\big)\,.$$

We intend to bound the right-hand side by some strictly negative number, which will imply the decreasingness of $\Delta_q T^{(u)} V$. Adding and subtracting $\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0)$, this becomes:

$$(\tilde{\Lambda} + \theta)\left(\Delta_q T^{(u)} V(n) - \Delta_q T^{(u)} V(n-1)\right) =$$
$$\left(\alpha - \mu + \frac{\Lambda}{N}\right)\left(\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0)\right) \tag{34}$$
$$+ \frac{\Lambda}{N}\left(\Delta_n V(n, 0) - \Delta_n V(n-1, 0) - (\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0))\right). \tag{35}$$

For (34), we use the assumptions $\Delta_n V(n-1, 0) - \Delta_n V(n-2) \geq \kappa > 0$ and $\alpha - \mu + \Lambda/N < 0$ to get:

$$\left(\alpha - \mu + \frac{\Lambda}{N}\right)\left(\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0)\right) < \kappa\left(\alpha - \mu + \frac{\Lambda}{N}\right).$$

For (35), we have: on the one hand with P1 a):

$$\Delta_n V(n-2, 0) \leq \frac{c_Q}{\alpha + \theta + \Lambda/N}, \qquad \Delta_n V(n, 0) \leq \frac{c_Q}{\alpha + \theta + \Lambda/N}$$

and on the other hand, with Picx and P1 b):

$$-\Delta_n V(n-1, 0) \leq -c_B\,.$$

In total,

$$\Delta_n V(n, 0) - \Delta_n V(n-1, 0) - (\Delta_n V(n-1, 0) - \Delta_n V(n-2, 0)) \leq 2\left(\frac{c_Q}{\alpha + \theta + \Lambda/N} - c_B\right)$$

and therefore:

$$(\tilde{\Lambda} + \theta)\left(\Delta_q T^{(u)} V(n) - \Delta_q T^{(u)} V(n-1)\right) < \kappa\left(\alpha - \mu + \frac{\Lambda}{N}\right) + \frac{2\Lambda}{N}\left(\frac{c_Q}{\alpha + \theta + \Lambda/N} - c_B\right).$$

The limit of the right-hand side when $N \to \infty$ is $\kappa(\alpha - \mu) < 0$. Therefore, there exists N_0 such that $\forall N \geq N_0$, this right-hand side is strictly negative. Summing up, for every such N, provided assumptions a) and b) hold true,

$$\Delta_q T^{(u)} V(n) - \Delta_q T^{(u)} V(n-1) < 0$$

for all $2 \leq n \leq N$, which proves $iii)$. □

From a system point of view, the presence of two dynamics can be interpreted but not explained. We first recall that since Assumption (28) holds then it is always more interesting to serve a customer than not to serve it. However, from the curves (as well as monotonicity) several trends can be observed. When the loss rate α is larger than the departure rate μ, then the more customers there are in the queue the less serving them is interesting. Indeed, the absolute value of the difference between $T_{AS}^{(u)}$ and $T_{NS}^{(u)}$ decreases. Furthermore, the sign of the difference does not change which means that the optimal decision does not change. Instead, when the loss rate is smaller than the service rate then the more customers there are in the queue the more the service is interesting.

This change of structural properties furthermore illustrates observations that can be made from the literature. In [18], the authors analyze a slightly different model that also includes impatience in service. It is proved that submodularity holds under the condition that the total departure rate of the server (service rate and impatience in service rate) is larger than the loss rate. Since our rate of impatience in service is null then we have a similar condition. At last, similar changes of dynamics can be found in [2] which aims at computing the optimal service rate in a queueing model with impatience. Indeed, their cost function is either convex or concave according to the relative values of the reneging rate and the service rate.

7.4 Limits of truncated non-smoothed models

We analyze here the boundary effects of truncation and smoothing on the structural properties of the value function. Example 5 seeks to illustrate how truncation breaks structural properties at the boundary while in Example 6 one examines all the possible combinations between truncation smoothing and shows that only the truncated smoothed version of this work preserves structural properties.

For the clarity of this part, wee call "*bounded by S*" a model in which the state space is finite and in which the maximum number of customers is S. In $x = (S, b)$ we jump toward the same state with a rate λ.

We call "*truncated by N*" a model in which the rate of the exponential distributions of arrivals or/and departure depends on the state and is null for all $x = (n, b)$ with $n \geq N$. The state space can be either finite or infinite, but when the state space is bounded by S, S should be larger than N.

We call "*smoothed in N*" a model truncated by N in which the rate of the exponential distributions of arrivals or/and departure depends on the state and decreases with respect to the number of customers such that it is equal to 0 in $x = (N, b)$.

We call uniformized a model in which the transition rate is the same whatever the state is.

Example 5 (Effect of the truncation)**.** *We consider first the Base model of Definition 2, with parameters $\alpha(n) = n\alpha$ and $\lambda(n) = \Lambda$ for all $n \geq 0$. The value function of the "no serve" policy in this case is given by* (33) *in Theorem 5.*

Consider now the truncated model in N in which $\lambda(n) = \Lambda$ when $0 \le n \le N$ and $\lambda(n) = 0$ otherwise. In a such model, the fixed point equation of the "no serve" policy becomes:

$$V(n,0) = \begin{cases} \dfrac{1}{\Lambda + n\alpha + \theta}\left[nc_Q + \Lambda V(n+1,0) + n\alpha V(n-1,0)\right] & \text{for } 0 \le n \le N \\[2ex] \dfrac{1}{n\alpha + \theta}\left[nc_Q + n\alpha V(n-1,0)\right] & \text{for } n > N\,. \end{cases}$$

Let $\check{V}(n,0)$ be a function solution of the fixed point equation defined by the two equations just above. Then, we have that $\check{V}(n,0) = \widehat{V}(n,0)$ for $n \le N$ and, when $n > N$, then it satisfies the recurrence

$$\check{V}(n,0) = \check{V}(n-1,0)\left(1 + \frac{\theta}{n\alpha}\right)^{-1} + c_Q\left(\alpha + \frac{\theta}{n}\right)^{-1}\,.$$

When n tends to infinity then the function $\check{V}(n,0)$ approaches the linear line of slope c_Q/α. This slope is different than $c_Q/(\alpha+\theta)$ which is the slope of $\widehat{V}(n,0)$. Moreover, this asymptotic slope c_Q/α is steeper than the slope $c_Q/(\alpha+\theta)$ of $\check{V}(n,0)$ when $n \le N$.

On the other hand, at $n = N+1$ we have

$$\begin{aligned} \check{V}(N+1,0) - \check{V}(N,0) &= \frac{(N+1)c_Q - \theta\check{V}(N,0)}{(N+1)\alpha + \theta} \\ &= \frac{c_Q}{\alpha+\theta}\frac{1}{(N+1)\alpha+\theta}\left((N+1)(\alpha+\theta) - (N\theta + \Lambda)\right) \\ &= \frac{c_Q}{\alpha+\theta}\,\frac{(N+1)\alpha + \theta - \Lambda}{(N+1)\alpha+\theta} \\ &< \frac{c_Q}{\alpha+\theta} \;=\; \check{V}(N,0) - \check{V}(N-1,0)\,. \end{aligned}$$

This shows that $\check{V}(n,0)$ is not convex in N while $\widehat{V}(n,0)$ is.

Therefore we show that the truncation modifies the behavior of the fixed point solution of a Bellman operator and in particular that this may break convexity properties.

Example 6. *We consider the parameters $\Lambda = 0.5$, $\alpha = 1.0$, $\mu = 5.0$ and $\theta = 1.5$. We also consider $c_B = 1.0$, $c_L = 2$ and $c_H = 2$ (yielding $c_Q = 4$). We used a bounded model since only in bounded models numerical computations can be done. The value of S is chosen to be equal to 75. Our experiments with larger value of S, namely $S = 750$, $S = 7500$, $S = 10000$ produced variations in the value functions smaller than 10^{-7} at any state $x = (n,0)$ with $n \le N$. These variations are sufficiently small for our claim and it seems unnecessary to try to get smaller values of gaps.*

We combine truncation, smoothing to build different models. Thus we study the following models:

	Bounded	Uniformized	Arrival	Impatience
Model 1	Bounded	No	Not truncated	No truncated
Model 2	Bounded	No	Truncated $N = 50$	No truncated
Model 3	Bounded	No	Truncated $N = 50$	Truncated $N = 50$
Model 4	Bounded	No	Not truncated	Truncated $N = 50$
Model 5	Bounded	No	Smoothed $N = 50$	Truncated $N = 50$
Model 6	Bounded	Yes	Not truncated	Truncated $N = 50$
Model 7	Bounded	Yes	Truncated $N = 50$	Truncated $N = 50$
Model 8	Bounded	Yes	Smoothed $N = 50$	Truncated $N = 50$

For each of these models, we numerically compute the value function and we observe some of their structural properties, namely, monotonicity and convexity, as defined in Section 5.1. The results are compiled in Table 1, in which we also indicate if the value iteration method finds the optimal policy.

	Monotonicity	Convexity	Concavity	Optimal policy
Model 1	Yes	$\forall n \leq 73$	no	yes
Model 2	Yes	$\forall n \leq 73$	no	yes
Model 3	Yes	$\forall n \leq 49$	no	yes
Model 4	Yes	$\forall n \leq 49$	$\forall n \geq 50$	no $\forall n \geq 55$
Model 5	Yes	$\forall n \leq 49$	$\forall n \geq 50$	no $\forall n \geq 55$
Model 6	Yes	$\forall n \leq 49$	$\forall n \geq 50$	yes
Model 7	Yes	$\forall n \leq 49$	$\forall n \geq 50$	yes
Model 8	Yes	$\forall n \leq 50$	$\forall n \geq 51$	yes

Table 1: Studies of structural Properties for different models

The results presented here show that except for Model 8 *(which is the one studied in this work) all the other ones experience border effects for the structural properties. Indeed, in bounded models, border effects are due to the change in the operator in state* S*: the term in* $\lambda(S)V(S+1,0)$ *is replaced by a term in* $\lambda(S)V(S,0)$*. In truncated models, border effects are due to the change in the operator in state* N*: the term in* $\lambda(N)V(N+1,0)$ *disappears. Smoothing allows to reduce the impact of this change in the operator and allows to keep structural properties.*

8 Conclusions

We present here a detailed treatment of a stochastic dynamic control problem with unbounded rates, coming from a queuing model with impatience. It allows us to prove the optimality of a policy that is either Always Serve customers or Never Serve customers. We also give a closed form formula of the value function when the Never Serve policy is optimal.

The results presented here are consistent with these ones of Aalto [1], concerning the problem of admission control in a batch service queue. The model of [1] further allows holding costs in service. When we set these costs to 0, the batch size to 1, and set in our model the impatience rate to 0, our results are the same. Indeed, adopting our notations and setting the batch size to 1, Aalto states that if $c_H/\theta \leq c_B$ then the No Service policy is optimal and instead, when $c_H/\theta \geq c_B$, a queue length threshold policy is optimal. If we replace α by 0 in Theorem 1, we obtain the same condition. Note however that our proof does not cover this limiting case.

The extension of our model to more complex (non-linear) holding costs appears to be possible by adapting the tools presented here. The extension to batch sizes larger than 1 as in [1] is not straightforward, since the value function is then not convex in x.

At last, this work can be seen as a step toward an unified treatment of the propagation result for structural properties in case of unbounded rates as suggested in [9], or for the special cases of models with impatience.

References

[1] S. Aalto. Optimal control of batch service queues with finite service capacity and linear holding costs. *Math Meth Oper Res*, 51:263–285, 2000.

[2] M. Armony, E. Plambeck, and S. Seshadri. Sensitivity of optimal capacity to customer impatience in an unobservable M/M/S queue (why you shouldn't shout at the DMV). *Manufacturing and Service Operations Management*, 11(1):19–32, 2009.

[3] O. Baron, O. Berman, and D. Perry. Continuous review inventory models for perishable items ordered in batches. *Mathematical Methods of Operations Research*, 72(2):217–247, 2010.

[4] S. Benjaafar, J.-P. Gayon, and S. Tepe. Optimal control of a production-inventory system with customer impatience. *Operations Research Letters*, 38:267–272, 2010.

[5] E. Berk and U. Gurler. Analysis of the (Q,r) inventory model for perishables with positive lead times and lost sales. *Operations Research*, 56(5):1238–124, 2008.

[6] P. P. Bhattacharya and A. Ephremides. Optimal scheduling with strict deadlines. *IEEE Trans. Automatic Control*, 34(7):721–728, July 1989.

[7] S. Bhulai, A. Brooms, and F. Spieksma. On structural properties of the value function for an unbounded jump Markov process with an application to a processor sharing retrial queue. *Queueing Systems*, 76(4):425–446, 2014.

[8] H. Blok and F. Spieksma. Countable state Markov decision processes with unbounded jump rates and discounted cost: Optimality equation and approximations. *Adv. Appl. Prob.*, 47:1088–1107, 2015.

[9] H. Blok and F. Spieksma. *Markov Decision Processes in Practice*, chapter Structures of Optimal Policies in MDPs with Unbounded Jumps: The State of Our Art, pages 131–186. Springer, 2017.

[10] R. Boucherie and N. van Dijk, editors. *Markov Decision Processes in Practice*. Springer, 2017.

[11] R. K. Deb and R. F. Serfozo. Optimal control of batch service queues. *Advances in Applied Probability*, 5(2):340–361, 1973.

[12] D. Down, G. Koole, and M. Lewis. Dynamic control of a single-server system with abandonments. *Queueing System*, pages 63–90, 2011.

[13] X. Guo and O. Hernández-Lerma. *Continuous-Time Markov Decision Processes – Theory and Applications*. Springer, 2009.

[14] J. Hasenbein and D. Perry. Introduction: queueing systems special issue on queueing systems with abandonments. *Queueing Systems*, 75:111–113, 2013.

[15] E. Hyon and A. Jean-Marie. Scheduling services in a queuing system with impatience and setup costs. *The Computer Journal*, 55(5):553–563, 2012.

[16] I. Karaesmen, A. Scheller-Wolf, and B. Deniz. *Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook*, chapter Managing Perishable and Aging Inventories: Review and Future Research Directions, pages 393–436. Springer, 2011.

[17] G. Koole. Monotonicity in Markov reward and decision chains: Theory and applications. *Foundation and Trends in Stochastic Systems*, 1(1), 2006.

[18] M. Larrañaga, U. Ayesta, and I. Verloop. Asymptotically optimal index policies for an abandonment queue with convex holding cost. *Queueing systems*, 2-3:99–169, 2015.

[19] M. Larrañaga, O. Boxma, R. Nuñez-Queija, and M. Squillante. Efficient content delivery in the presence of impatient jobs. In *Proceedings of 27th ITC*, pages 73–81, 2015.

[20] A. Movaghar. Optimal control of parallel queues with impatient customers. *Performance Evaluation*, 60:327–343, 2005.

[21] K. P. Papadaki and W. B. Powell. Exploiting structure in adaptative dynamic programming algorithms for a stochastic batch service problem. *European Journal of Operational Research*, 142:108–127, 2002.

[22] M. Puterman. *Markov Decision Processes Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[23] S. Vakilinia, X. Zhang, and D. Qiu. Analysis and optimization of big-data stream processing. In *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016.

[24] A. Ward and S. Kumar. Asymptotically optimal admission control with impatient customers. *Mathematics of Operations Research*, 33(1):167–202, 2008.

[25] H. Weiss. Optimal ordering policies for continuous review perishable inventory models. *Operations Research*, 28(2):365–374, 1980.

A Proof of Proposition 1

The proof is based on Theorem 4.10, p. 60 of Guo and Hernández-Lerma [13]. We first state this theorem and the assumptions it relies on. We next apply it to our situation.

A.1 Theorem and assumptions

According to Remark 4.9, p. 63 *op. cit.*, Theorem 4.10 applies when Assumption 4.12 (p. 64) is satisfied. A second condition to be satisfied is Assumption 2.2 (p. 13).

In the context of the theorem, the state space is denoted with S, the action space available in state i is $\mathcal{A}_i$, the cost function is $c(i,a)$ and the transition rates are $q(j|i,a)$. The cost function must be nonnegative, which is the case in our application, see Definition 1. The conditions state as follows.

Assumption 2 (Assumption 2.2. from [13])**.** *There exists a function $w : S \to \mathbb{R}$, $w \geq 1$, and constants $c_0 \neq 0$, $b_0 \geq 0$ and $L_0 \geq 0$ such that:*

(a) $\sum_j q(j|i,a)w(j) \leq c_0 w(i) + b_0$ *for all* (i,a)*;*

(b) $\sup_a q(i,a) \leq L_0 w(i)$ *for all* i*.*

Assumption 3 (Assumption 4.12. from [13])**.** *(a)* $\mathcal{A}_i$ *is compact for each* $i \in S$*.*

(b) For all $i, j \in S$*, the functions* $c(i,a)$ *and* $q(j|i,a)$ *are continuous on* $\mathcal{A}_i$*.*

(c) There exists $\hat{f} \in F$ *such that* $V^{\hat{f}} < \infty$*, and* $\sum_j q(j|i,a)V^{\hat{f}}(j)$ *is continuous on* $\mathcal{A}_i$ *for all* i*.*

Here, as in Section 3, F is the class of state-feedback stationary controls and $V^{\hat{f}}(j)$ is the expected cost of control $\hat{f}$ from initial state j.

Then, the result is:

Theorem 6 ([13], Theorem 4.10)**.** *Let Assumptions 2 and 3 hold. Then for all* $x \in S$*,*

$$V^*(x) \;=\; \min_{a \in \mathcal{A}_x} \left\{ \frac{c(x,q)}{q(x,a)+\theta} + \frac{1}{q(x,a)+\theta} \sum_{y \in S; y \neq x} q(y|x,a)\, V^*(y) \right\}, \tag{36}$$

and any $f \in F$ *such that* $f(x)$ *belongs to the "arg min" in (36) is optimal.*

A.2 Application to Proposition 1

We claim that under the assumptions of Proposition 1, Assumptions 2 and 3 are satisfied by the formal model.

Observe first that, thanks to the finiteness of the sets $\mathcal{A}_x$, all compactness and continuity requirements of Assumption 3 are indeed satisfied. We just need to check statement (c), that is, the existence of a Markov policy with a finite cost. This is brought by the following Lemma. Recall the definition of the "no service" policy π_{NS} in Definition 3.

Lemma 14. *Under the assumptions of Proposition 1, the policy* π_{NS} *yields a finite value* V^{NS}*.*

Proof. By assumption, the arrival rate $\lambda(\cdot)$ is bounded above by some Λ. Also by assumption and from Definition 1, the running cost $c(x,a)$, $x=(n,b)$, is bounded by a polynomial in n. Indeed,

$$c(x,a) \le (q(x,a)+\theta)c_B+k(n-a) \le (\Lambda+\max(\alpha(n),\alpha(n-1))+\mu+\theta)c_B+\max(k(n),k(n-1)).$$

Since $\alpha(.)$ and $k(.)$ are bounded by polynomials, $c(x,a) \le K(n)$ for some polynomial K, assumed to be increasing without loss of generality.

Consider the queue with constant arrival rate Λ, no admission in service and no impatience (in other words, with $\alpha(n)=0$ for all n), and running cost $c^U(x,a)=K(n)$. In this queue (labeled with "U"), no departure ever occurs and $x^U(t) =_d x^U(0)$ + Poisson(Λt). By a coupling argument (including the coupling of service times if a customer is in service at $t=0$), it is clear that $x(t) \le_{st} x^U(t)$ and since $K(\cdot)$ is increasing, it follows that $V^{NS} \le V^U$. It is sufficient to prove the finiteness of V^U. If $0=T_0<T_1<\dots$ denotes a Poisson process with rate Λ, we have:

$$\begin{aligned}
V^U(x) &= \mathbb{E}\left[\int_0^\infty e^{-\theta t} K(x^u(t))\,\mathrm{d}t\right] \\
&= \mathbb{E}\left[\sum_{\ell=0}^\infty \int_{T_\ell}^{T_{\ell+1}} e^{-\theta t} K(x(0)+\ell)\,\mathrm{d}t\right] \\
&= \mathbb{E}\left[\sum_{\ell=0}^\infty K(x(0)+\ell) e^{-\theta T_\ell}\,\frac{1-e^{-\theta(T_{\ell+1}-T_\ell)}}{\theta}\right] \\
&= \sum_{\ell=0}^\infty K(x(0)+\ell)\left(\frac{\Lambda}{\Lambda+\theta}\right)^\ell \frac{1}{\Lambda+\theta}\,.
\end{aligned}$$

This quantity is indeed finite for all $\theta>0$ since $K(\cdot)$ is a polynomial. □

We now turn to Assumption 2. The following Lemma is useful for checking this assumption and other ones. In the statement, $Q(a)$ denotes the rate matrix made of rates $q(y|x,a)$, $(x,y)\in\mathcal{X}\times\mathcal{X}$.

Lemma 15. *For any $\gamma>0$, there exists some $\varepsilon>0$ such that the function w_ε: $\mathcal{X}\to\mathbb{R}$, defined as $w_\varepsilon(n,b)=e^{\varepsilon n}$, satisfies: $Q(a)w_\varepsilon \le \gamma w_\varepsilon$ for all a.*

Proof. The fact that γ satisfies $Q(a)w_\varepsilon \le \gamma w_\varepsilon$ is granted if

$$\sup_{n,b} \frac{(Q(a)w_\varepsilon)(n,b)}{w_\varepsilon(n,b)} \;<\; \gamma\,.$$

We therefore proceed with the computation of this supremum.

Let us start with the case $x=(n,1)$, in which case $\mathcal{A}_x=\{0\}$. We have (see Section 2.2):

$$\begin{aligned}
(Q(a)w_\varepsilon)(n,1) &= \lambda(n)w_\varepsilon(n+1,1) \;+\; \alpha(n)w_\varepsilon(n-1,1) \;+\; \mu w_\varepsilon(n,0) \\
&\quad -\; (\lambda(n)+\alpha(n)+\mu)w_\varepsilon(n,1) \\
&= e^{\varepsilon n}\left(\lambda(n)(e^\varepsilon-1)+\alpha(n)(e^{-\varepsilon}-1)\right),
\end{aligned}$$

then

$$\frac{(Q(a)w_\varepsilon)(n,1)}{w_\varepsilon(n,1)} \;=\; \lambda(n)(e^\varepsilon-1)+\alpha(n)(e^{-\varepsilon}-1) \;\le\; \Lambda(e^\varepsilon-1)$$

since $\lambda(n)$ is bounded, by assumption. Next, consider a state $(n, 0)$ such that $a = 0$, that is, no service is started. Then,

$$\begin{aligned}(Q(a)w_\varepsilon)(n,0) &= \lambda(n)w_\varepsilon(n+1,0) \;+\; \alpha(n)w_\varepsilon(n-1,0) \;-\; (\lambda(n)+\alpha(n))w_\varepsilon(n,0) \\ &= e^{\varepsilon n}\left(\lambda(n)(e^{\varepsilon}-1)+\alpha(n)(e^{-\varepsilon}-1)\right),\end{aligned}$$

then

$$\frac{(Q(a)w_\varepsilon)(n,0)}{w_\varepsilon(n,0)} \;=\; \lambda(n)(e^{\varepsilon}-1)+\alpha(n)(e^{-\varepsilon}-1) \;\leq\; \Lambda(e^{\varepsilon}-1)\,.$$

Finally, if $a = 1$ (and $n \geq 1$), then:

$$\begin{aligned}(Q(a)w_\varepsilon)(n,0) &= \lambda(n-1)w_\varepsilon(n,1) \;+\; \alpha(n-1)w_\varepsilon(n-2,1) \;+\; \mu w_\varepsilon(n-1,0) \\ &\quad -\; (\lambda(n-1)+\alpha(n-1)+\mu)w_\varepsilon(n,0) \\ &= e^{\varepsilon n}\left(\alpha(n-1)(e^{-2\varepsilon}-1)+\mu(e^{-\varepsilon}-1)\right),\end{aligned}$$

and

$$\frac{(Q(a)w_\varepsilon)(n,0)}{w_\varepsilon(n,0)} \;=\; \alpha(n-1)(e^{-2\varepsilon}-1)+\mu(e^{-\varepsilon}-1) \;\leq\; 0.$$

Summing up all three cases, we find that for all a and all $x \in \mathcal{X}$,

$$\sup_{x\in\mathcal{X}} \frac{(Q(a)w_\varepsilon)(x)}{w_\varepsilon(x)} \;\leq\; \Lambda(e^{\varepsilon}-1)$$

and clearly this can be made smaller than any $\gamma > 0$ by a proper choice of ε, namely, $\varepsilon \leq \log(1+\gamma/\Lambda)$. This proves the statement. □

As a consequence of Lemma 15, Assumption 2 (a) holds with $c_0 = \gamma$ and $b_0 = 0$, where γ can be any non-negative number. The existence of L_0 as in Assumption 2 (b) is granted if

$$\sup_{(n,b)} \frac{\sup_a q((n,b),a)}{w(n,b)} \;<\; \infty. \tag{37}$$

Given the values of $q((n, b), a)$ in Section 2.2 and the polynomial growth of $\alpha(\cdot)$, this property also holds.

Therefore, Assumption 2 holds, and Theorem 6 applies. This proves Proposition 1.

B Proof of Proposition 2

The proof relies on Theorem 4.2 of [8]. We first state the theorem and the assumptions it relies on. We then apply it to our situation.

B.1 Theorem and assumptions

We use Theorem 4.2 of [8]. It applies to a collection of parametrized Markov Reward processes $\{X(a); a \in \mathcal{A}\}$ on some state space S, whose infinitesimal generator is denoted with $Q(a)$, whose running cost function is denoted with $c(x, a)$ and whose value function is the expected total discounted cost with discount factor α.[10] It holds under the conditions summarized in Assumption 4 below.[11]

Assumption 4. *a) (Assumption 2.1 from [8]) For each $a \in \mathcal{A}$, $X(a)$ is a minimal, standard, stable Markov process with right-continuous sample paths and with conservative q-matrix $Q(a)$;*

b) (Assumption 3.1 from [8]) The set A is a locally compact topological space, i.e. every point in A has a compact neighborhood;

c) (Assumption 3.2 (i) from [8]) the functions $a \mapsto q_{xy}(a)$ are continuous for every x, y in S;

d) (Assumption 3.2 (ii) from [8]) there exists a function V and a real number γ such that for all $a \in \mathcal{A}$,

$$Q(a)V \;\leq\; \gamma V;$$

e) (Assumption 3.2 (iii) from [8]) there exists a function W and a real number θ with, for all $a \in A$,

$$Q(a)W \;\leq\; \theta W$$

and an increasing sequence of finite sets $\{K_n\}$ with $\lim_n K_n = S$ and

$$\lim_{n\to\infty} \inf\{W(x)/V(x); x \in K_n\} \;=\; +\infty.$$

f) (Assumption 4.1 (i) from [8]) the functions $a \mapsto c(x, a)$ are continuous for every x in S;

g) (Assumption 4.1 (ii) from [8]) there is a finite real constant c_V such that $\sup_{x,a} |c(x, a)|/V(x) \leq c_V$ (in other words: $\sup_{x,a} |c(x, a)|/V(x) < \infty$);

h) (Assumption 4.1 (iii) from [8]) the discount factor α is such that $\alpha > \gamma$;

i) (Assumption 4.2 from [8]) There exist compact metric sets $\mathcal{A}_x$, $x \in S$, such that $\mathcal{A} = \prod_x \mathcal{A}_x$ and is equipped with the product topology, and for any $a, a' \in \mathcal{A}$, $x \in S$, such that $a_x = a'_x$, it holds that $(Q(a))_x = (Q(a'))_x$ and $c(x, a) = c(x, a')$.

If Assumption 4 $c)$ holds, then it is possible to define $\ell^\infty(S, V)$, the space of V-bounded functions on S.

Theorem 7 ([8], Theorem 4.2)**.** *Suppose that Assumption 4 holds. Suppose moreover than $\sup_a q(x, a) < \infty$ for all $x \in S$. Then the equation*

$$\alpha f(i) \;=\; \inf_{a\in A(i)} \left\{ c(i, a) + \sum_{j\in\mathcal{X}} q(j|i, a) f(j) \right\}$$

has a unique solution $v^\alpha \in \ell^\infty(S, V)$ and the infimum is a minimum. Every policy a^* that achieves the minimum is optimal.

[10] This parameter a is not to be confused with the individual actions a of the formal model in Section 2.2. The matrix $Q(a)$ is not to be confused with that of Appendix A.

[11] Note that Theorem 4.2 in [8] is stated without the explicit Assumption 3.1 *op. cit.* However, the proof does require Theorem 4.1, which holds under this assumption.

B.2 Application to Proposition 2

We now check that all conditions of Assumption 4 are satisfied for the family of models defined in Definition 1, parametrized by the Markov (feedback) policy $\pi \in F$. We therefore set $\mathcal{A} = F$ in the application of Theorem 7. This is a discrete set, hence locally compact. Requirement b) is then satisfied. All Markov processes we consider have a finite number of transitions out of each state, with finite rates. Condition a) is therefore satisfied.

As observed earlier, the action space $\mathcal{A}$ does satisfy condition i) with $\mathcal{A}_x = \{0, 1\}$ or $\{0\}$. These sets are discrete and finite so that all compactness and continuity requirements c) and f) are satisfied.

The existence of a function V as in d), with a constant $\gamma < \alpha$ as in h), is a consequence of Lemma 15: for every γ, we can chose some $\varepsilon > 0$ and $V(n, b) = e^{\varepsilon n}$. The existence of a function W as in e) also derives from this lemma. It suffices to take $W(n, b) = e^{\varepsilon' n}$ for some $\varepsilon' > \varepsilon$, and the sets K_n can simply be chosen as $K_n = \{(i, b), i \leq n, b \in \{0, 1\}\}$.

With this choice for the function V, the boundedness of $c(x, a)/V(x)$ is a consequence of the polynomial growth assumed on $k(\cdot)$, since $V(\cdot)$ has an exponential growth. Requirement g) is therefore satisfied.

All requirements of Assumption 4 are therefore checked. There remains to check that $\sup_a q(x, a) < \infty$ for all $x \in S$. This is so because $\mathcal{A}_x$ is a finite set for all x.

C Convergence of models

We use Theorem 5.1 from [8]. It applies to a collection of parametrized Markov processes $\{X(N, \delta); (N, \delta) \in \mathcal{N} \times F\}$, on some discrete, denumerable state space S, where F is the set of admissible stationary deterministic policies (*i.e.* feedback policies, see Section 3.1) and $\mathcal{N} = \mathbb{N} \setminus \{0\} \cup \{\infty\}$.

Assumption 5. *Let* $A = \mathcal{N} \times F$*:*

a)-h) Same assumptions as in Assumption 4 for this set A*;*

i) for every fixed N*, the family of processes* $\{X(N, \delta); \delta \in F\}$ *has the "product property" (that is: Assumption 4* h*) for the set* $A = F$*, and for every* N*).*

Theorem 8 ([8], Theorem 5.1)**.** *Consider a collection of parametrized Markov processes* $\{X(N, \delta); (N, \delta) \in \mathcal{N} \times F\}$ *and cost function* $c : \mathcal{N} \times F \to \mathbb{R}$.

Suppose that Assumption 5 holds. Let v_N^α *be the value function for the MDP* $\{X(N, \delta)\}$ *and* δ_N^* *an optimal policy. Then the following hold:*

(i) $\lim_{N \to \infty} v_N^\alpha = v_\infty^\alpha$*;*

(ii) any limit point of $(\delta_N^*)_{N \in \mathcal{N}}$ *is optimal for* $X(\infty, \delta)$.

In the application of this result, observe that the set $\mathcal{N}$ is compact when equipped with the metric:

$$d_{\mathcal{N}}(n, m) \;=\; \left| \frac{1}{n} - \frac{1}{m} \right|, \quad m, n \neq \infty, \qquad d_{\mathcal{N}}(n, \infty) \;=\; \frac{1}{n}, \quad n \neq \infty.$$

Therefore, if F is locally compact, then $A = \mathcal{N} \times \mathcal{F}$ is locally compact also, which validates Assumption 5b).

Contents

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399