Proposal for an ARC INRIA
OCOQS: Optimal threshold policies in COntrolled Queuing Systems

1 Title and executive summary

Title: Optimal Threshold Policies In Controlled Queuing Systems

Acronym: OCOQS

Summary: The research domain of this proposition is the optimal control of stochastic processes, with applications to the control of networks and manufacturing systems.

The principal aim is to widen the set of mathematical techniques that can be used to prove that optimal policies are of *threshold* type, thereby widening the set of classes of models that can be effectively solved exactly or numerically handled in practice.

To that end, the collaborative action will bring together specialists: of discrete-event systems and stochastic orders on the one hand, and of queueing, networking and manufacturing applications on the other hand. By analyzing several concrete problems which cannot presently be handled by known methods, it is hoped that new proof techniques, presumably based on stochastic ordering concepts, will emerge.

2 Head of the ARC

The head of this Collaborative Research Action is Ana Bušić, member of the TREC team of INRIA Paris-Rocquencourt (joint team with ENS Paris).

3 Teams and Members

The teams and researchers directly involved in this ARC are:

- Ana Bušić, Junior Research Scientist (CR2), team TREC of INRIA Paris-Rocquencourt and ENS Paris (Paris antenna).
  Her research fields are Markov chains and Markov reward processes, stability in queuing systems, performance evaluation of communication networks.
  `http://www.di.ens.fr/~busic/`.
- Alain Jean-Marie, Senior Research Scientist (DR2), team MAESTRO of INRIA Sophia-Antipolis Méditerranée (Montpellier antenna).
  His research interests include queueing theory, optimal control and game theories.
  `http://www.lirmm.fr/~ajm/`.

- Emmanuel Hyon, Assistant Professor at University of Paris Ouest Nanterre La Defense and associate researcher in the RO team of LIP6.
  His research interests are related with stochastic optimization and especially optimal control in queuing systems or networks.
  `http://www-sysdef.lip6.fr/~hyon/`.
- Ingrid Vliegen, Assistant Professor at the School of Management and Governance of the University of Twente (UT), the Netherlands.
  Her research fields are optimization of health care processes and manufacturing systems.
  `http://www.utwente.nl/mb/ompl/staff/Vliegen/`.

All the participants have an expertise in the analysis of stochastic Discrete Events Dynamics Systems (SDEDS) mainly using the framework of queuing theory. They are currently interested in the optimal control of such systems. The analysis techniques used by these participants are applicable to a quite large variety of problems. These techniques are the Markov Decision Process (MDP) framework for MAESTRO, RO and UT and the Stochastic Orders for TREC and MAESTRO.

From the point of view of specific models and applications, some teams (MAESTRO, TREC, RO) are implied in the theoretical analysis of networks (seen as dynamic systems) and some of them (RO, UT) are interested in inventory and supply chain systems.

The participants to this project come from two different research communities: communication networks and manufacturing systems. Both communities use queueing models but they have developed different methods. For the "networking" community, analysis methods consist mainly in using stochastic orders while in the "manufacturing systems" community, they are mainly based on propagation of structural properties.

This proposal is a consequence of existing collaborations between either MAESTRO and RO [20, 11] or TREC and UT [9, 43] and this would be the first common participation to some research project for this group.

Finally, let us mention that if the natural application fields concern network and queueing theory, nevertheless our initial motivations directly follow from industrial contracts on which we have worked in the past. An optimal planning determination problem for a satellite launcher system implies RO and MAESTRO teams [11] while during her PhD thesis [42], I. Vliegen worked with ASML, within the context of Service Logistics Forum (SLF).

4 Context, motivation and main objectives

4.1 General context

The main objectives of this ARC are related with the optimal control of stochastic systems. The framework of Markov Decision Processes (MDPs) provides a well-established theory for this question. From a practical point of view however, actually computing optimal control rules (or "policies") is difficult due to the (in)famous "curse of dimensionality". In many cases however, the optimal policy has a very simple *structure*: by just comparing the value of some element of the current state to a fixed value (a *threshold*), one of two actions can be selected. Such policies are called *threshold policies* (or *control limit policies*) and finding the best one in their family is a much easier practical problem. Using a *structured* analysis of

MDPs (see Section 7.1), it is possible to prove that an optimal policy is of threshold type[1] without actually computing it. This will be the issue tackled in this ARC.

4.2 Scientific objectives

The *structured analysis* framework for MDPs is now well established but suffers limitations (see again Section 7.1).

Our higher-level aims in this collaborative research action are:

1. Extend (if possible) the list of elementary discrete-event constructions that lead to structural properties, in order to enlarge the family of models that can be handled with this approach.

2. Render structural MDP analysis more systematic.

   Indeed, the application of the methodology often requires a tedious case-by-case analysis, which is model-dependent.

Our belief is that the systematic use of stochastic ordering concepts may be the key to such goals.

In order to demonstrate that this idea is valid, we plan to start with its application to concrete queueing models. Several publications exhibiting special structural properties have appeared in the area of manufacturing systems. The use of such a structural analysis of MDP models in Networking is more recent. We think that there are opportunities for transferring knowledge and methods from one application domain to the other.

Our shorter-term objectives for this ARC are then:

1. Solve (that is, compute the optimal policy for) several models representative of open problems in the application of MDPs both to Manufacturing or Networking problems.[2]

2. Try to adapt the existing structural properties which are used in different problems of management science to network problems.

4.3 Para-scientific objectives

In addition to these focused scientific objectives, we seek through the proposed ARC to:

1. generally increase the expertise of the participants, taking advantage of the variety of techniques and applications used by the different teams, and also by inviting experts outside of the consortium for seminars and short visits;

2. identify a network of researchers interested in launching a large-scale, international action focused on structured analysis of MDPs.

Possible followups to this project will depend on the scope of the network identified: either some ANR project, or some bi-lateral Van Gogh action, up to some larger European project if appropriate.

[1] We also include policies that have two thresholds such as "(s, S)" policies appearing in stock management problems.

[2] Work has already begun on these models, as explained in Section 5.2.

5 Scientific program

Our approach to handle the scientific objectives above will be to start with concrete models, on which we have already started to work, and for which we have identified: a) that the optimal policy is possibly of threshold type, b) that the known structural analysis methods do not apply, c) possible solution techniques to apply. By systematically considering stochastic order properties of these controlled dynamical systems, we hope to be able to identify common features and general principles.

We provide in this section a general description of the mathematical tools we plan to develop (Section 5.1), then a description of the models we shall study (Section 5.2). The methodological background is reviewed in Section 7.

5.1 Methodology

The standard approach to study Markov Decision Processes (MDP) concentrates on results that are obtained by inductively proving properties of the dynamic programming value function. To simplify the proofs, the dynamics of the system is decomposed in several operators that are studied separately (see Section 7 for a more detailed presentation). This approach has its limits when the system is complex:

- the individual operators do not necessarily propagate inductively the structural properties (such as monotonicity or convexity),
- on the global level (when we consider all operators together), the dynamic programming value function is not necessarily convex at all times,
- this approach is model dependent and tedious,
- in addition, there is a lack of some unified method for actually computating threshold values.

In what follows, we give research directions that will be considered in this ARC proposal to cope with these limits.

Composition of operators. When operators which describe the effects of events or controls do not satisfy any structural characteristics individually, one can try to combine them into more complex event operators. The issue consists in finding the appropriate composition, such that it corresponds with the model and the dynamics of the model and such that the global operators satisfy some structural feature. Furthermore, this structural feature should be related with the properties of the policy one wants to show. When this idea has already been successfully applied in the literature [2], it led to compact proofs. However for other constructions, the solutions proposed in the literature remain model-driven.

Weaker properties. When the standard properties do not hold then weaker properties could still be exhibited. This amounts to trying to soften the structural properties while still keeping the desired characteristics of the policy. The issue is to find the appropriate property which is compatible with the dynamics of the system and that implies the feature of the optimal policy. The idea originates from works in inventory problems (see later). Hence, these works have shown that convexity is only a sufficient condition to show that a policy is of threshold type. The notion of K-convexity [16] is used for that purpose, but the possible transposition of K-convexity to other models remains unknown.

Structural properties and stochastic orders. Some structural properties seem to be naturally related to stochastic orders of random variables and Markov processes. Some obvious examples are preservation of increasing property and stochastic monotonicity using strong stochastic order, or preservation of convexity and stochastic monotonicity using convex order. The link with stochastic orderings for some weaker notions of convexity is less obvious.

The aim is to identify the links between the structural properties and stochastic orders. Both approaches seem to have their advantages. Stochastic orders allow an unified treatment for a whole family of value functions and are less tedious to work with. On the other hand, the results obtained using stochastic orders are much stronger and in some cases they do not hold, while it is still possible to have some structural properties for a given value function.

We expect that stochastic orders can also help to better understand the composition rules that can be applied when the individual operators do not propagate the structural properties.

Threshold computation. All the previous points addressed the question how to prove some structural properties that imply a threshold optimal policy. This then reduces significantly the space of all candidate policies. However, this does not give any mean to select the optimal one in the threshold policy class and the set of all threshold policies can be too big for an exhaustive search.

In some rare cases it is possible to obtain analytic expressions for the optimal threshold. This can be done by fixing two different values for the threshold and then comparing the two Markov reward processes, using either dynamic programming approach or stochastic orders. We expect that the latter can give less tedious proofs.

If the threshold cannot be computed exactly, we are left with numerical methods that suffer from the state space explosion problem. Thus, we want to bound the optimal threshold value. Instead of analyzing directly an original, unsolvable model, we can design, using Stochastic Comparison ideas, a new model satisfying the following two properties:

- the modified model provides bounds for the given value function;
- it is tractable.

This new model can then be used to obtain the bounds for the optimal threshold.

5.2 Problems addressed

We detail below some problems which are addressed by our models. Our intention is to solve them during the duration of this ARC.

Control of queues with impatience

In [20], we have proposed a queueing model with batch service in discrete time. Customers which arrives are stored in a queue in which they wait to be serviced. Customers must be served before some time limit (deadline) otherwise they leave the system. 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 (called “losses”) induce some costs. The purpose is to decide when to serve (when to control) the customers so as to minimize these costs.

The problem with batches of size one has been solved. We shall study the extension to batches of arbitrary size, as well as other extensions: considering other costs functions such as costs depending on the throughput of the system, and systems where there are different types of customers.

A polling system with a feedback loop

We consider a specific type of polling system with a feedback loop between the queues, i.e. when a customer is served at one queue, there is a possibility that this customer joins another queue. An application of this is a hospital, where patients are seen by a doctor first at the outpatient clinic. Afterwards, a customer might be seen by the same doctor in the operating theater, which might be followed by a follow-up appointment again at the outpatient clinic. In this case, the doctor is the server that switches between several queues (customers in different stages of the process). Whenever the doctor is in the operating room (OR), he performs surgery on a batch of patients (i.e., the OR is not opened for just one patient). The objective is to decide which type of client to serve as to maximize the throughput, with the additional constraints that the waiting time for a patient needs to be below some threshold. The latter can be modelled using impatience.

A stock rationing system

We consider a single-product stock rationing system with multiple demand classes and lost sales. Upon a demand of a type k we can decide to satisfy the demand or to reject the demand which then has a certain cost c_k that depend on the type of demand. The stock can be modeled as a single class queue representing the items on stock, satisfied demands can be seen as service and replenishment as arrivals into the queue. The aim is to decide when to satisfy a demand, depending on the state of the system and demand type. We expect that the known results (see Section 7.2) on the optimality of the threshold policies can be extended to the non-exponential (e.g. hyperexponential or phase type) replenishment times.

Demand substitution

Inventory models can be seen as queueing models where the stock is represented by available servers, and items in replenishment are modeled as busy servers. Queueing jobs represent a backlog situation, whereas the Erlang loss model can be used to model a lost sales situation. We extend this basic model with a possibility of substitution. Thus, whenever an item is out of stock (or has limited stock), there is an option to give a customer another (usually more expensive) substitute. The goal is to decide when to use the substitute in order to maximize the revenue.

6 Expected results

The interaction between members of this ARC should create some scientific exchanges and methodological approach debates and comparisons.

Main advances expected are quite theoretical and methodological results. These are:

- the extension of the list of structural properties which propagate by dynamic programming operators;
- the development of a new approach based on stochastic orders. This one is completely new and should allow to fill the gaps left by the standard approach in its present state.

Concerning the applications, the advances should be:

- the resolution of the models presented in this document;
- the extension of the set of problems treated by the theoretical methods above.

From a para-scientific point of view, the participants gain a better cross knowledge of both fields: structural approaches and stochastic orders.

7 State of the art

7.1 Markov decision processes and structured analysis of DES

The theory of Markov Decision Processes (MDPs) is now well established. We make here a quick review, with the purpose of introducing concepts and identifying the limitations to which we plan to remedy.

We consider a system which behavior is driven by random events (called transitions), and which can be controlled to some extent. Denote with x_n is the state of the system at the epoch of the n^{th} transition and by q_n the control. The evolution of the state from x_n to x_{n+1} is described by some dynamic operator R, through the recurrence equation: $x_{n+1} = R(x_n, q_n)$. A *policy* is a sequence of decision rules $\pi = (d_0, d_1, \ldots)$, each decision rule d_i mapping some information set to some control. Under a given policy, the evolution of the system generates a random sequence of states x_n and decisions q_n. To each policy is associated a cost:[3]

$$v_\theta^\pi(x) \;=\; \mathbb{E}_x^\pi \left[\sum_{n=0}^{\infty} \theta^n \, c(x_n, q_n) \right] \;,$$

for some discount factor θ, and where $c(x, q)$ expresses the instantaneous cost incurred by taking decision q when the state is x. The question is to find the policy which optimizes (minimizes or maximizes) the value function. The value of this optimal policy (also called the *value function* of the MDP problem) is the unique solution to the Bellman equation, which is expressed in functional form as:

$$v = \min_q (Tv)(x, q) \;,$$

with T the dynamic programming operator, itself expressed as:

$$(Tv)(x, q) = c(x, q) \;+\; \theta \, \mathbb{E}\left[v\left(R(x, q) \right) \right] \;.$$

[3]We consider here the average discounted cost to simplify this presentation.

Structured policies framework

The framework of *structured* analysis of MDP problems has been exposed in the books of Glasserman and Yao [18], Puterman [32] and Bertsekas [5], to name only three of them. It consists in showing that the optimal policy is "structured" in the sense that it has a special form (for example increasing, decreasing...). Together with the notion of structured policies, comes a notion of structured value functions. Both notions are adapted to each other.

The technique consists in propagating such structural properties through the operator T defined above. Loosely described: if v is in some class of functions, say V^{σ}, then Tv has certain properties, and $\min_q Tv(\cdot, q)$ is also in V^{σ}. In the limit, the optimal value function is in V^{σ} as well, and the structure of the optimal policy is deduced.

The success of this approach is linked to properties enjoyed by the value function V and the "Q-value" function TV. Common properties found in the literature are monotonicity, convexity/concavity, and sub/supermodularity. Some of these properties can be related to *stochastic ordering* properties of the operator R.

Lately, Koole has proposed in [22] an inventory of elementary constructions, based on an identification of types of events in the SDEDS, which lend themselves to a structural analysis. He identifies at the same time weaker yet useful structural properties, such as 2-convexity.

Limitations

If this framework has been successfully applied to many models, there are also many cases where its application is not straightforward.

It often happens that the optimal policy exhibits some good properties (as far as numerical solutions can tell) while the dynamic operators do not check usual structural properties at all: one says that the structural properties does not propagate. This phenomenon has been identified in different counter examples: systems with losses in [22, 20], multi items systems (each type of customers induces different reward values) in [10].

In some other problems, it is the value function which does not satisfy any useful feature while the optimal properties does. This makes the proofs much more involved. Examples of this include [29] or see Section 4.2 in [5] in which standard techniques do not hold.

At last, we point out that this kind of structured analysis suffers to be model dependent. Hence, very often, a tedious case by case analysis must be performed to obtain structural results. Unfortunately, they must be performed again each times a new problem is considered.

7.2 Applications of MDPs

Problems related to Optimal Control using Markov Decision Processes have been largely studied in the literature. Indeed, they cover a large spectrum of applications in the fields of communication networks or in Operations Research in general.

Control of queuing systems. Controlled queuing models, deterministic as well as stochastic, have been largely studied in the literature since their application fields are numerous. To quote just a few, it can be mentioned networking (see [14] for a general presentation and [1] for problems solved with MDPs) or resource allocation (see [21] and references therein).

We shall be interested in the control of queues with impatience: the phenomenon of impatience, associated with deadlines has become a major interest in several fields of networking: cellular communication networks [3], call centers [23], best route determination [26, 35]).

However, the literature does not usually consider impatient customers or losses; for instance, the problem of optimally controlling a batch server in a queue without impatience has been addressed in [29].

This kind of problems with perishable items equally appear in management science. They belong to the class of Network Revenue Management [6]. They are mainly studied for yield management problems (seat reservation of airline companies) or reservations problems (see [30] for an overview and see [11] for a satellite launcher system).

Polling systems. A typical polling system consists of a number of queues, attended by a single server in a fixed order, with a non-negative switching time between the queues. There is a lot of literature on polling systems with a wide range of applications in communication, production, transportation, maintenance systems and healthcare. Surveys on polling systems and their applications may be found in [34] and [41].

Inventory systems. For a single product stock rationing system, with a *single* server and multiple demand classes and lost sales (first studied by Veinott Jr [40]), it is known that the optimal policy is a basestock threshold policy (see e.g. Ha [19], De Véricourt, Karaesmen and Dallery [12]). For ample repair capacity (infinite server), a basestock policy is known to be suboptimal. However, if the basestock level is fixed, then the optimal policy is still a threshold policy in some cases. For exponentially distributed repair times and ample repair capacity, a threshold policy is proven to be optimal by Miller [25], and Van Jaarsveld and Dekker [39]. The problem of optimization within the class of threshold policies is adressed in Dekker et al. [13] and Kranenburg and Van Houtum [24]. For a single repair server with Erlang-k distributed repair times, the optimality of a threshold policy is proven by Ha [19] (for lost sales) and Gayon et al. [17] (for backordering).

Examples of substitution in inventory systems can be found in manufacturing, where expensive tool kits are used as back-up for several separate tools [44], in call centers, where employees with multiple skills act as a back-up for colleagues that are only trained on one skill, and in retail, where customers decide to take another item when the desired item is out of stock [28].

7.3 New structural properties and stochastic orders

Non standard structural properties. When standard comparisons are useless, it appears that more complicated properties have to be found. These issues were addressed in very few works.

Altman and Koole [2] identify problems for which submodularity does not hold. However, the operator exhibits a weaker property which propagates and which insures monotonicity of control. In order to do that, the authors are able to express the global operator under the form of a decomposition of simple operators.

Working on value function properties in inventory problems, [31] details how some special threshold policies (the (s, S)-policies) can be shown to be optimal while the value function is not convex (as the standard approach requires). Instead, the value function has its own special property: the K-convexity property see [16].

Stochastic orders. The stochastic order framework is a well known way to get several performance values (or bounds on these values) of a stochastic system [27, 33]. This can be done without computing exactly the full distribution.

Stochastic orders appear sometimes in specific arguments in proofs for optimal control problems. For instance, we have used one such argument in [20]. However, to the best of our knowledge, the literature does not feature a the single, systematic framework for optimal control questions, based on stochastic orders.

Bounding Markov chains and Markov reward processes. Stochastic comparison equally allows to construct *simpler models* by modifying states and transitions of the model. These simpler models can be solved using classical methods and thus, *provide bounds* for steady-state or transient distributions. Several algorithms have been developed in the literature that are based on comparison of random variables and the monotonicity property of Markov chains [36, 15, 4, 8, 7].

While stochastic comparison gives results for an arbitrary chain and for a rather general family of rewards, the Markov reward approach [37] is more suitable when we want to obtain bounds for a specific system and a given reward. A systematic method to construct bounds by using this approach is given in [38].

In the existing collaboration between TREC and UT, a new method has been developed that extends this approach to redirecting selected sets of transitions (which is related to preservation of convexity and supermodularity properties) [9]. It was successfully applied to prove new monotonicity results and bounds for assemble-to-order systems.

References

[1] E. Altman. Applications of Markov Decision Processes in communication networks. In E.A. Feinberg and A. Shwartz, editors, *Handbook of Markov Decision Processes*, chapter 16. Kluwer, 2002.

[2] E. Altman and G. Koole. On submodular value functions and complex dynamic programming. *Stochastic Models*, 14:1051–1072, 1998.

[3] J. R. Artalejo and V. Pla. On the impact of customer balking, impatience and retrials in telecommunication systems. *Computers and Mathematics with Applications*, 57(2):217–229, 2009.

[4] M. Ben Mamoun, A. Busic, and N. Pekergin. Generalized class C Markov chains and computation of closed-form bounding distributions. *Probability in the Engineering and Informational Sciences*, 21(2):235–260, 2007.

[5] Dimitri Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, second edition, 2001.

[6] D. Bertsimas and I. Popescu. Revenue management in a dynamic network environment. *Transportation Science*, 37:257–277, 2003.

[7] A. Busic. *Comparaison stochastique de modèles markoviens : une approche algorithmique et ses applications en fiabilité et en évaluation de performance.* PhD thesis, Université de Versailles Saint-Quentin-en-Yvelines, 2007.

[8] A. Busic and J.-M. Fourneau. Bounds based on lumpable matrices for partially ordered state space. In *SMCtools '06: Proceeding from the 2006 workshop on Tools for solving structured Markov chains, workshop of 1st International Conference Valuetools'06*. ACM Press, 2006.

[9] A. Bušić, I. Vliegen, and A. Scheller-Wolf. Comparing Markov chains: Combining aggregation and precedence relations applied to sets of states, 2009. Rapport de recherche. `http://hal.archives-ouvertes.fr/hal-00361430/fr/`. Submitted.

[10] E. B. Çil, E. L. Örmeci, and F. Karaesmen. Structural results on a batch acceptance problem for capacitated queues. *Mathematical Methods of Operations Research*, 66:263–274, 2007.

[11] P. Chrétienne, E. Hyon, and A. Jean-Marie. Assessing performance requirements for launch systems with dual launch capability - part I: Simulator and modeling. ESA Reference: Purchase Order noP351003, juillet 2005.

[12] F. de Véricourt, F. Karaesmen, and Y. Dallery. Optimal stock allocation for a capacitated supply system. *Management Science*, 48(11):1486–1501, 2003.

[13] R. Dekker, R.M. Hill, M.J. Kleijn, and R.H. Teunter. On the $(S-1, S)$ lost sales inventory model with priority demand classes. *Naval Research Logistics*, 49(6), 2002.

[14] A. Ephremides and S. Verdu. Control and optimization methods in communication network problems. *IEEE Transactions on Automatic Control*, 34:930–942, 1989.

[15] J.-M. Fourneau and N. Pekergin. An algorithmic approach to stochastic bounds. In *Performance Evaluation of Complex Systems: Techniques and Tools, Performance 2002, Tutorial Lectures*, volume 2459 of *LNCS*, pages 64–88. Springer, 2002.

[16] G. Gallego and S. Sethi. K-convexity in R^n. *Journal of Optimization Theory and Applications*, 127(1), 2005.

[17] J.P. Gayon, F. de Véricourt, and F. Karaesmen. Stock rationing in an M/E r/1 multi-class make-to-stock queue with backorders. *IIE Transactions*, 41(12):1096–1109, 2009.

[18] P. Glasserman and D. Yao. *Monotone Structure in Discrete-Event Systems*. Wiley, 1994.

[19] A.Y. Ha. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43(8):1093–1103, 1997.

[20] E. Hyon and A. Jean-Marie. Scheduling in a queuing system with impatience and setup costs. In *ISCIS 2010, 25th International Symposium on Computer and Information Sciences*, London, September 2010.

[21] A. J. Kleywegt and J. D. Papastavrou. The dynamic and stochastic knapsack problem. *Operations Research*, 46:17–35, 1998.

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

[23] G. Koole and A. Mandelbaum. Queueing models of call centers: An introduction. *Annals of Operations Research*, 113:41–59, 2002.

[24] A.A. Kranenburg and G.J. Van Houtum. Cost optimization in the $(S-1, S)$ lost sales inventory model with multiple demand classes. *Operations Research Letters*, 35(4):493–502, 2007.

[25] B.L. Miller. A queueing reward system with several customer classes. *Management Science*, 16(3):234–245, 1969.

[26] D. Mitra and Q. Wang. Stochastic traffic engineering, with applications to network revenue management. In *Infocom*. IEEE, 2003.

[27] A. Muller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley, New York, NY, 2002.

[28] S. Netessine and N. Rudi. Centralized and competitive inventory models with demand substitution. *Operations Research*, 2003.

[29] K. Papadaki and W. Powell. Exploiting structure in adaptive dynamic programming algorithms for a stochastic batch service problem. *EJOR*, 142:108–127, 2002. `http://citeseer.ist.psu.edu/595069.html`.

[30] J. D. Papastavrou, S. Rajagopalan, and A. J. Kleywegt. The dynamic and stochastic knapsack problem with deadlines. *Management Science*, 42(12):1706–1718, 1996.

[31] E.L. Porteus. *Foundations of Stochastic Inventory Theory*. Stanford Univ. Press, 2002.

[32] M.L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc. New York, NY, USA, 1994.

[33] M. Shaked and J. G. Shantikumar. *Stochastic Orders and their Applications*. Academic Press, San Diego, CA, 1994.

[34] H. Takagi. Analysis and application of polling models. In C. Lindemann G. Haring and M. Reiser, editors, *Performance Evaluation: Origins and Directions*, number 1769 in Lecture Notes in Computer Science, pages 423–442. Springer, Berlin, 2000.

[35] H. Topaloglu. Using lagrangian relaxation to compute capacity-dependent bid prices in network revenue management. *Operations Research*, 57:637–649, 2009.

[36] L. Truffet. Reduction technique for discrete time Markov chains on totally ordered state space using stochastic comparisons. *Journal of Applied Probability*, 37(3):795–806, 2000.

[37] N.M. van Dijk. Bounds and error bounds for queueing networks. *Annals of Operations Research*, 79:295–319, 1998.

[38] G. J. van Houtum, W. H. M. Zijm, I. J. B. F. Adan, and J. Wessels. Bounds for performance characteristics: a systematic approach via cost structures. *Commun. Statist. - Stochastic Models*, 14(1&2):205–224, 1998.

[39] W. van Jaarsveld and R. Dekker. Finding optimal policies in the $(S-1, S)$ lost sales inventory model with multiple demand classes. *Report Econometric Institute*, EI 2009-14, 2009.

[40] A.F. Veinott Jr. Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. *Operations Research*, 13(5):761–778, 1965.

[41] V.M. Vishnevskii and O.V. Semenova. Mathematical methods to study the polling systems. *Autom. Remote Control*, 2:173–220, 2006.

[42] I. Vliegen. *Integrated planning for service tools and spare parts for capital goods*. PhD thesis, Eindhoven University of Technology, 2009.

[43] I.M.H. Vliegen, A. Bušić, A.A. Scheller-Wolf, and G.J. van Houtum. Optimization of stock levels for service tool inventory, 2010. In preparation.

[44] I.M.H. Vliegen, P.A.M. Kleingeld, and G.J. van Houtum. Separate tools or tool kits: an exploratory study of engineers' preferences. *International Journal of Production Economics*, 125(1):173–184, 2010.