Probability in the Engineering and Informational Sciences, **18**, 2004, 13–31. Printed in the U.S.A.

CONVEXITY IN TANDEM QUEUES

GER KOOLE
Department of Mathematics
Vrije Universiteit Amsterdam
The Netherlands
E-mail: koole@cs.vu.nl

We derive convexity results and related properties for the value functions of tandem queuing systems. The results for standard multiserver queues are new. For completeness, we also prove and generalize existing results on tandems of controllable queues. The results can be used to compare queuing systems. This is done for systems with and without batch arrivals and for systems with different numbers of on–off sources.

1. INTRODUCTION

During the past two decades, the study of controlled queuing systems has been a fertile area of research. One well-researched notion is that optimal controls are often monotone in the state of the system. For example, under certain assumptions about costs, the optimal choice of a server's rate turns out to be an increase in the size of the queue that is served. This type of result is usually obtained by proving monotonocity properties of value functions of queuing systems. However, these monotonicity results can also be useful to derive other types of results, such as comparing the performance of different systems. As an example, consider the standard single-server $M/M/1$ queue. It is well known that the dynamic programming value function is a convex function of the number of customers in queue for many standard criteria. This is the basic step to show that for any additional controlled arrivals stream, customers should be admitted only up to a certain threshold level (e.g., see [10]). On the other hand, convexity is crucial in comparing uncontrolled single-server queues with different numbers of on–off arrival processes (see [9]), and for comparing tandem systems consisting of $\cdot/G/s$ queues with their fluid limits [2]. This shows that it is equally interesting to prove monotonicity results for controlled queuing

© 2004 Cambridge University Press 0269-9648/04 $16.00

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

systems as for uncontrolled queuing systems. Conceptually, there is no difference in the treatment of the two types of system.

Let us consider the different classes of system for which monotonicity properties have been derived. For one- and two-dimensional systems, in our opinion, a fairly complete picture is given in [8]. The results for systems of higher dimension fall apart in first-order and second-order results. The first-order results compare value functions in different states (e.g., optimality of the $c\mu$ rule in [5], optimality of shortest queue routing in [7]); the second-order results deal with properties related to convexity. The only result known in the last class is that of [16] (and its generalizations in [3]). It considers a tandem (or a cycle) of queues where each queue has a controlled departure rate. For suitably chosen objective functions, this can be translated to results for single-server queues. However, this cannot be done for multiserver queues. For this type of queuing system, there are no known results.

In this article, we present monotonicity results for tandems of multiserver queues. For the sake of completeness and to facilitate comparison, we also present results for controlled queues.

Central in our approach is dynamic programming (dp), by propagating certain properties related to convexity through the dynamic programming value function. The concept of *event-based dp* (as introduced in [8]) allows us to treat classes of value functions at once and to present the results in a concise and clear way. It consists of cutting up the value function into smaller parts—the *event operators*. Each operator corresponds to a (physical) event in the system, such as arrivals or departures from a certain queue. For the direct costs and the random environment, there are also operators to enable us to compose the value function entirely of event operators. We formally introduce event-based dp and the operators that we consider in this article in Section 2.

After having formulated the subject of the study—the operators that constitute the value function—we have to formulate the properties of which we will inductively show that the value function satisfies them. These properties are formulated in terms of inequalities that are mostly related to convexity. There are two types of second-order property. The first is *multimodularity*, a property of functions on the lattice introduced in [6]. The second is a property that is known as *directional convexity* (see [14]). Section 3 is devoted to the study and comparison of multimodularity, directional convexity, and related properties. Characterizations are given that simplify the proofs in later sections.

The main results are derived in Section 4. For each class of functions (characterized by the second-order properties), those operators are selected that are closed under the set of functions. From this, it follows that value functions with these operators as building blocks are elements of the same class of functions. The first class to be considered are multimodular functions, and results equivalent to those obtained in [16] for tandem queues are derived in the more general framework of event-based dp. The operators are characterized by the fact that the service rate needs to be controllable. By restricting the class of multimodular functions to those which have also a first-order property, we can allow for standard single-server queues.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

However, we argue that multimodularity is not the right class for multiserver queues. Therefore, we move to the class of directional convex functions (with an additional first-order property) and we show that the multiserver operators are closed under this set. These results are valid for tandem systems of arbitrary dimensions; the result was only known for the considerably simpler two-dimensional case [8].

Applications that are based on multimodularity or directional convexity of the value function are related to the comparison of queuing systems. We give examples of two types of comparison result. The first type is concerned with systems with different types of arrival process, but with the same average number of arrivals per time unit. The first example compares a system with Poisson arrivals with the same system, except for the arrival process, which is twice as slow, but every arrival consists of a batch of size two. It is shown that the batch arrivals give a worse performance for all performance measures for which the value function is directional convex. The second example generalizes the result of [9] to tandem systems. In [9], systems with identical on–off sources are compared, and it is shown that a system with K on–off sources behaves better than one with one on–off source (but with a K times higher arrival rate).

The second type of comparison result is concerned with showing that the performance of the tandem system is convex in one of the parameters. As an example, we show that this is indeed the case for the rate at which customers arrive from the exterior into one of the queues, given that the value function itself is directional convex. An application of this type of result is in load balancing between parallel (tandem) systems. Suppose that customers arrive at a service facility and they have to be routed in a Bernoulli fashion to a parallel processing unit, each consisting of a tandem system. Convexity of the performance of each tandem queue in the Bernoulli parameter makes the problem of minimizing total performance a convex programming problem, which is easy to solve. Of course, similar results can be derived for other system parameters. This type of result can be seen as a generalization of results in [11] to tandem systems. All of these applications can be found in Section 5.

The results in this article parallel, to some extent, the results obtained in [12, Chap. 3]. They were obtained independently and used completely different methods. From a methodological point of view, both Meester and I find that both methods merit publication.

2. MODELS

The method we use to derive our results is inductively showing certain properties of the dynamic programming (dp) value function. A general tool for simplifying and generalizing this task is *event-based dp*, as introduced in [8]. In that article, it is also applied to certain classes of models, which are in essence one- or two-dimensional. Our current interest is in higher-dimensional models. As event-based dp simplifies the current analysis as well, we start with describing the main ideas behind it. After that, we analyze the current class of systems.

Event-based dp is based on two crucial ideas. One is to split up the dp value function into smaller components—the *event operators*. The other is to give the set

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

of functions that is to be propagated a central place, instead of the model itself. For a general description of event-based dp, see [8]. We restrict ourselves right away to the type of model that is studied in this article.

We take $\mathbb{N}_0^{m+1}$ as the state space (with $\mathbb{N}_0 = \{0,1,\ldots\}$). For $x = (x_0,\ldots,x_m) \in \mathbb{N}_0^{m+1}$, the component x_i, $i > 0$, represents the number of customers at queue i. The variable x_0 is the state of the *environment*. Let $\mathcal{V}$ be the class of functions from $\mathbb{N}_0^{m+1}$ to $\mathbb{R}$. We will define the event operators—the building blocks of the value function. The operators that we are using in this article are of two types: controlled and uncontrolled; for example, we consider both a regular server that is active as long as there are customers waiting to be served and a server of which the service rate can be controlled. One of the benefits of event-based dp is that we consider both types of operator in the same framework, which makes clear what difference in model assumptions is necessary to deal with different operators.

In giving names to the operators, we use *C* for "controlled," *A* for "arrivals," *D* for "departures," *TS* for "tandem server," and *M* for "multiple." The operators $(\mathcal{V} \to \mathcal{V})$ are, for all $f \in \mathcal{V}$:

$$
\begin{aligned}
T_{A(i)} f(x) &= f(x + e_i),\\
T_{CA(i)} f(x) &= \min\{f(x), c + f(x + e_i)\},\\
T_{D(i)} f(x) &= f((x - e_i)^+),\\
T_{CD(i)} f(x) &= \begin{cases} \min\{c + f(x - e_i), f(x)\} & \text{if } x_i > 0\\ f(x) & \text{otherwise,}\end{cases}\\
T_{TS(i)} f(x) &= \begin{cases} f(x - e_i + e_{i+1}) & \text{if } x_i > 0\\ f(x) & \text{otherwise,}\end{cases}\\
T_{CTS(i)} f(x) &= \begin{cases} \min\{c + f(x - e_i + e_{i+1}), f(x)\} & \text{if } x_i > 0\\ f(x) & \text{otherwise,}\end{cases}\\
T_{MD(i)} f(x) &= \begin{cases} x_i f(x - e_i) + (S - x_i) f(x) & \text{if } x_i < S\\ S f(x - e_i) & \text{otherwise,}\end{cases}\\
T_{MTS(i)} f(x) &= \begin{cases} x_i f(x - e_i + e_{i+1}) + (S - x_i) f(x) & \text{if } x_i < S\\ S f(x - e_i + e_{i+1}) & \text{otherwise.}\end{cases}
\end{aligned}
$$

Of course, the operators incorporating one or more tandem servers are only defined for $i = 1,\ldots,m-1$; the other operators are defined for $i = 1,\ldots,m$. Furthermore, for ease of notation, we used c in every operator that involves costs for some action. This c is arbitrary and can be different for each operator.

The coefficients of f in the operators can be interpreted as probabilities, except for the operators with multiple servers: There the coefficients sum to S. We could divide by S to obtain probabilities, but this division could also be part of the envi-

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

ronment operator to be introduced next. To keep the notation as simple as possible, we assume the latter.

In addition to these operators dealing with the movements of customers in the system, we introduce an operator for the direct costs and one that determines how the environment (i.e., x_0) changes state and influences the transitions:

$$T_{\text{costs}} f(x) = C(x) + \alpha f(x)$$

for some cost function C and discount factor $\alpha \geq 0$,

$$T_{\text{env}}(f_1, \ldots, f_l)(x) = \sum_{y \in \mathbb{N}_0} \lambda(x_0, y) \sum_{j=1}^{l} q^j(x_0, y) f_j(x^*),$$

where x^* is equal to x with the zeroth component replaced by y.

We see that $T_{\text{env}}: \mathcal{V}^l \to \mathcal{V}$. To see how these operators can constitute the value function V_n, it is best to consider an example that models a simple tandem system of $M/M/1$ queues (the server at queue i has rate μ_i) with Poisson arrivals (at rate λ) and some arbitrary cost function:

$$V_{n+1}(x) = T_{\text{costs}}[T_{\text{env}}(T_{A(1)}(V_n), T_{TS(1)}(V_n), \ldots, T_{TS(m-1)}(V_n), T_{D(m)}(V_n))](x), \quad \textbf{(1)}$$

with the parameters of T_{env} chosen such that it models the well-known concept of uniformization (e.g., see [13, Sec. 11.5]), by taking $T_{\text{env}}(f_1, \ldots, f_{m+1})(x) = \sum_{j=1}^{m+1} p(j) f_j(x)$ with $p(1) = \lambda/(\lambda + \sum_i \mu_i)$ and $p(i) = \mu_{i-1}/(\lambda + \sum_j \mu_j)$, $i = 2, \ldots, m+1$. That V_n is indeed equal to the value function of the continuous-time tandem server model can be easily seen by writing out (1).

The operator T_{env}, however, can serve many more purposes. We see that if the environment moves from x_0 to y, which occurs with probability $\lambda(x_0, y)$ (but, due to the uniformization, we can also see this as a transition rate), then with probability $q^j(x_0, y)$, an event of type j occurs. This is exactly the definition of a *Markov arrival stream* [4]. A Markov arrival stream can best be interpreted as a Markov arrival process with additional marks indicating the type of event. It is shown in [4] that any marked point process can be approximated arbitrarily close with a Markov arrival stream. This shows the versatility in modeling capacities of T_{env}. A relatively simple example of the use of T_{env} can be found in Section 5, where it is used to model parallel on–off processes.

Describing formally how V_{n+1} can be constructed from V_n is done in [8]. We will not repeat it here because from (1), the main ideas will be clear. With this construction, many different systems with different criteria can be modeled. The main classes of models are those in discrete time and continuous time (modeled with uniformization). The standard criteria (finite horizon, infinite horizon, discounted, and average costs) can be modeled as well. (Note that for the infinite horizon models there are existence questions to be addressed. We will not deal with them here, but will concentrate on the properties of V_n for finite n.) However, more general systems can be addressed as well, which are not time-homogeneous, which have salvage costs, and so forth.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

The next step is formulating sets of functions $\mathcal{F} \subset \mathcal{V}$ such that if $V_n \in \mathcal{F}$, then so is V_{n+1} for particular choices of V_n. For these sets, we show, for some or all of the operators T introduced above, that if $f \in \mathcal{F}$, then also $Tf \in \mathcal{F}$. From this, it readily follows that if $V_0 \in \mathcal{F}$ and $C \in \mathcal{F}$, then also $V_n \in \mathcal{F}$ for all n. The sets $\mathcal{F}$ should be chosen such that some useful results can be derived from them. The different sets $\mathcal{F}$ and their comparison are the subject of the next section.

Remark 2.1: In [16,8] $T_{CTS(i)}$ is generalized to a continuously controlled operator of the form $\min_{\mu\in[0,1]}\{c_\mu + \mu f(x - e_i + e_{i+1}) + (1-\mu)f(x)\}$, with zero as the only possible action if $x_i = 0$. The results of the next sections remain the same under this generalization; we have decided not to include it, as it complicates the notation.

3. INEQUALITIES

The sets of functions that we consider are defined by inequalities. Remember that our state space is $(m+1)$ dimensional. However, we will pose no restrictions on the environment; thus the inequalities are essentially m dimensional. The inequalities are both of first and second order. The two basic second-order properties are *directional convexity* and *multimodularity*. We start with directional convexity (dc).

DEFINITION 3.1: *A function $f \in \mathcal{V}$ is called directional convex if, for all $x, y, z \in \mathbb{N}_0^{m+1}$, $y_0 = z_0 = 0$,*

$$f(x+y) + f(x+z) \leq f(x) + f(x+y+z),$$

which we write as $\mathrm{DC}(y,z)$.

Later, when studying queuing models, it will become clear that Definition 3.1 is difficult to use in our proofs. Therefore, we derive an alternative characterization of directional convexity. To do so, we define the following inequalities for some function $f \in \mathcal{V}$, for all $x \in \mathbb{N}_0^{m+1}$, $i, j \in \{1, \ldots, m\}$, $i \neq j$:

- $\mathrm{CC}(i)$: $2f(x+e_i) \leq f(x) + f(x+2e_i)$
- $\mathrm{Super}(i,j)$: $f(x+e_i) + f(x+e_j) \leq f(x) + f(x+e_i+e_j)$.

The inequalities are easily interpreted: CC stands for componentwise convexity and $\mathrm{Super}(i,j)$ is known as *supermodularity*. Note that $\mathrm{CC}(i)$ is equal to $\mathrm{Super}(i,i)$.

All functions satisfying a certain inequality A for all x are denoted by $\mathcal{F}(A)$. We also write $\mathcal{F}(A,B) = \mathcal{F}(A) \cap \mathcal{F}(B)$. The set of all dc functions is denoted by $\mathcal{F}(\mathrm{DC})$.

The characterization of directional convexity is as follows (see also [14]; for completeness, we also supply a proof).

LEMMA 3.2: $\mathcal{F}(\mathrm{CC}(i), i = 1, \ldots, m; \mathrm{Super}(i,j), i, j = 1, \ldots, m, i \neq j) = \mathcal{F}(\mathrm{DC})$.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

PROOF: The inclusion $\supset$ follows from the definition of directional convexity: taking $y = z = e_i$ gives $\mathrm{CC}(i)$ and taking $y = e_i$ and $z = e_j$ gives $\mathrm{Super}(i,j)$. The inclusion $\subset$ follows by induction to $\sum_i (y_i + z_i)$. If $\sum_i (y_i + z_i) = 2$, then $\mathrm{DC}(y,z)$ is either $\mathrm{CC}(i)$ for some i or $\mathrm{Super}(i,j)$ for some i and j. Assume that $\sum_i (y_i + z_i) > 2$ and suppose that $\sum_i y_i > 1$. Then, for some i, y can be written as $y = e_i + u$ with $u \geq 0$. Now $\mathrm{DC}(y,z)$ is the sum of

$$f(x + e_i + u) + f(x + e_i + z) \leq f(x + e_i) + f(x + e_i + u + z)$$

and

$$f(x + e_i) + f(x + z) \leq f(x) + f(x + e_i + z).$$

Both inequalities hold by induction. ■

Lemma 3.2 greatly simplifies the proofs in the next section: Instead of having to prove $\mathrm{DC}(y,z)$ for all possible y and z, we can restrict to y and z being unit vectors.

Next, we define multimodularity. It was introduced in Hajek [6] for the study of assignment rules to a queue, its relation to monotonicity results for tandem queues was recognized in [15, top of p. 295]. First, define $d_0 = e_1$, $d_i = -e_i + e_{i+1}$, $i = 1, \ldots, m-1$, and $d_m = -e_m$. Let $\mathcal{D} = \{d_0, \ldots, d_m\}$.

DEFINITION 3.3: *A function $f \in \mathcal{V}$ is called multimodular (mm) if for all $x \in \mathbb{Z}^{m+1}$, $v, w \in \mathcal{D}$, $v \neq w$, such that $x + v, x + w \in \mathbb{N}_0^{m+1}$*

$$f(x) + f(x + v + w) \leq f(x + v) + f(x + w),$$

which we write as $\mathrm{MM}(v,w)$.

Note that f is only defined on $\mathbb{N}_0^{m+1}$. The definition makes sense because it follows from $x + v, x + w \in \mathbb{N}_0^{m+1}$ that also $x, x + v + w \in \mathbb{N}_0^{m+1}$.

The usefulness of multimodularity lies in the fact that the basic transitions given in the operators are elements of $\mathcal{D}$. In the next lemma, we derive a more general characterization of multimodularity, which will help us in comparing multimodularity with directional convexity.

LEMMA 3.4: *A function f is mm if and only if*

$$f(x) + f(x + v_1 + \cdots + v_k) \leq f(x + v_1 + \cdots + v_l) + f(x + v_{l+1} + \cdots + v_k) \quad \textbf{(2)}$$

for $\{v_1, \ldots, v_k\} \in \mathcal{D}$, $0 < l < k$, and $v_i \neq v_j$ for $i \neq j$.

PROOF: The "if" part of the proof is trivial; the "only if" part follows by induction to k. For $k = 2$, (2) is equivalent to mm. Consider some $k > 2$ and assume that (2) holds up to $k - 1$. Furthermore, assume that $l + 1 < k$. This does not restrict generality; by reordering $v_1, \ldots, v_k$ in descending order in the case $l + 1 = k$, we find $l + 1 < k$ because $k > 2$. Equation (2) is equal to the sum of

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

$$f(x) + f(x + v_1 + \cdots + v_{k-1}) \leq f(x + v_1 + \cdots + v_l) + f(x + v_{l+1} + \cdots + v_{k-1})$$

and

$$f(x + v_{l+1} + \cdots + v_{k-1}) + f(x + v_1 + \cdots + v_k) \leq f(x + v_1 + \cdots + v_{k-1}) + f(x + v_{l+1} + \cdots + v_k).$$

Both inequalities hold by induction, the second in the state $x + v_{l+1} + \cdots + v_{k-1}$. ■

With $\mathcal{F}(\mathrm{MM})$, we denote the set of mm functions.

An inequality that will prove to be useful is *superconvexity*, introduced in [8] for the control of various two-dimensional queuing models. It is defined for all $i, j = 1, \ldots, m$, $i \neq j$, as

- $\mathrm{SuperC}(i,j)$: $f(x + e_i) + f(x + e_i + e_j) \leq f(x + 2e_i) + f(x + e_j)$.

For $m = 2$, it is readily seen that $\mathcal{F}(\mathrm{MM}) = \mathcal{F}(\mathrm{Super}(1,2), \mathrm{SuperC}(1,2), \mathrm{SuperC}(2,1))$. For general m, we have the following.

LEMMA 3.5: *$\mathcal{F}(\mathrm{MM}) \subset \mathcal{F}(\mathrm{SuperC}(i,j))$ for all $i \neq j$.*

PROOF: If $i < j$, then $\mathrm{SuperC}(i,j)$ is equal to (2) applied in $x + e_i$ with $d_0 + \cdots + d_{i-1} = e_i$ and $d_i + \cdots + d_{j-1} = -e_i + e_j$. If $i > j$, then it is (2) applied in $x + e_i + e_j$ with $d_j + \cdots + d_{i-1} = e_i - e_j$ and $d_i + \cdots + d_m = -e_i$. ■

Next, we compare the sets of dc and mm functions.

LEMMA 3.6: $\mathcal{F}(\mathrm{MM}) \subsetneq \mathcal{F}(\mathrm{DC})$.

PROOF: We first show $\subset$. Because of Lemma 3.2, it suffices to show for $f \in \mathcal{F}(\mathrm{MM})$ that $\mathrm{CC}(i)$ and $\mathrm{Super}(i,j)$ holds for all i, j. We do this using Lemma 3.4. We have $e_i = d_0 + \cdots + d_{i-1}$ and $-e_i = d_i + \cdots + d_m$. Equation (2) in $x + e_i$ gives $\mathrm{CC}(i)$. To show $\mathrm{Super}(i,j)$, assume (without restricting generality) that $i < j$. In addition to $e_i = d_0 + \cdots + d_{i-1}$ we need $-e_j = d_j + \cdots + d_m$ and $e_i - e_j = d_0 + \cdots + d_{i-1} + d_j + \cdots + d_m$. Equation (2) in $y = x + e_j$ is equivalent to $\mathrm{Super}(i,j)$.

We give a function on $\mathbb{N}_0^2$ that is dc but that does not satisfy $\mathrm{SuperC}(1,2)$. Because of Lemma 3.5, this function is not mm. This function is $f(x) = x_1 x_2$. Verifying $\mathrm{CC}(1)$, $\mathrm{CC}(2)$, and $\mathrm{Super}(1,2)$ in all x is simple. It is also easily seen that $\mathrm{SuperC}(1,2)$ does not hold for any x. ■

Remark 3.7: So far, we discussed the second-order property, directional convexity, and compared it to multimodularity. Another natural question is to compare both properties with the traditional convexity on the lattice, defined by the inequality $2f(x) \leq f(x - y) + f(x + y)$ that has to hold for each $y \in \mathbb{Z}^m$ and x such that $x - y$, $x + y \in \mathbb{N}_0^{m+1}$. In [1, Thm 2.2], it is shown that mm implies convexity; it is also easily seen that the function $f(x) = x_1 x_2$ is not convex (although we just showed in the

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

proof of Lemma 3.6 that x_1x_2 is dc). On the other hand, $g(x) = \max\{x_1, x_2\}$ is convex ([1], Counterexample 2.1), but g is not supermodular (e.g., Super(1,2) in $x = 0$ does not hold); thus, it is neither dc. In conclusion, if we denote with $\mathcal{F}(C)$ the set of convex functions $\mathcal{F}(\mathrm{MM}) \subset \mathcal{F}(C), \mathcal{F}(\mathrm{DC})$, but there is no such ordering between $\mathcal{F}(C)$ and $\mathcal{F}(\mathrm{DC})$.

This closes the discussion of second-order properties. We continue with first-order properties. Define the following inequalities for some function $f \in \mathcal{V}$, for all $x \in \mathbb{N}_0^{m+1}$, $i \in \{1, \ldots, m\}$:

- $I(i)$: $f(x) \leq f(x + e_i)$
- $F(i)$: $f(x + e_{i+1}) \leq f(x + e_i)$

Inequality $I(i)$ simply states that f is increasing (in the nonstrict sense) in i. $F(i)$ means that it is advantageous to have customers further down the line. Note that $F(i)$ is only defined for $i < m$. We introduce the notation $\mathcal{F}(I) = \mathcal{F}(I(i), i = 1, \ldots, m)$ and $\mathcal{F}(F) = \mathcal{F}(F(i), i = 1, \ldots, m-1)$.

LEMMA 3.8: *If $i < j$, then $\mathcal{F}(F(i), \ldots, F(j-1), I(j)) \subset \mathcal{F}(I(i))$ and, thus, $\mathcal{F}(F(1), \ldots, F(m-1), I(m)) = \mathcal{F}(I, F)$.*

PROOF: The proof simply consists of adding the inequalities that constitute the smaller (left hand) set in each possible state x. ■

4. RESULTS

In this section, we first discuss the operators that can be controlled. The suitable set of functions for these operators is $\mathcal{F}(MM)$. Later, we consider the other operators and functions. First, we assume that the value function is mm and we consider the consequences for the operators. After that, we prove that the value function is indeed mm for certain operators. The monotonicity results for the controlled operators shown in the proof are equal to those in [16] for controlled tandem systems. Our interest is mainly in the structure of the value functions. Note, furthermore, the generalizations with respect to the results of [16] due to event-based dp and the inclusion of the environment.

THEOREM 4.1: *For all $f \in \mathcal{F}(\mathrm{MM})$,*

$$T_{A(1)} f, \ldots, T_{A(m)} f, T_{CA(1)} f, T_{CTS(1)} f, \ldots, T_{CTS(m-1)} f, T_{CD(m)} f \in \mathcal{F}(\mathrm{MM}),$$

and if $C \in \mathcal{F}(\mathrm{MM})$ and $f_1, \ldots, f_l \in \mathcal{F}(\mathrm{MM})$,

$$T_{\mathrm{costs}} f, T_{\mathrm{env}}(f_1, \ldots, f_l) \in \mathcal{F}(\mathrm{MM}).$$

PROOF: Assume that $f \in \mathcal{F}(\mathrm{MM})$. Consider $T_{A(i)}$ for some i. We have to show that $T_{A(i)} f(x) + T_{A(i)} f(x + v + w) \leq T_{A(i)} f(x + v) + T_{A(i)} f(x + w)$ for all $v, w \in \mathcal{D}$. This is equivalent to $f(x + e_i) + f(x + e_i + v + w) \leq f(x + e_i + v) + f(x + e_i + w)$, which holds by induction. To facilitate the notation in the rest of this proof and other

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

proofs as well, we add the state to the inequality; thus, we have to show that $\mathrm{MM}(v,w)(x+e_i)$.

Now, we continue by considering $T_{CA(1)}, T_{CTS(1)}, \ldots, T_{CTS(m-1)}, T_{CD(m)}$, all at the same time. Let us introduce some further notation. Each of these operators T_C has the form $T_C f(x) = \min\{c + f(x+y), f(x)\}$, with $y \in \mathcal{D}$. We first prove the following property of the minimizers. The optimizing actions of $T_C f$ are such that the following hold:

- If $c + f(x+y) \leq f(x)$ for some x, then $c + f(x+d+y) \leq f(x+d)$ for any $d \in \mathcal{D}$ such that $d \neq y$.
- If $f(x) \leq c + f(x+y)$, then $f(x+y) \leq c + f(x+2y)$.

The first expression follows directly because mm is equivalent to $f(x) - f(x+y) \leq f(x+d) - f(x+d+y)$. For the second expression, we use Lemma 3.4 to get

$$f(x+y) + f\left(x+y+\sum_{v\in\mathcal{D}} v\right) \leq f(x+2y) + f\left(x+y+\sum_{v\in\mathcal{D}, v\neq y} v\right).$$

Because $\sum_{v\in\mathcal{D}} v = 0$ and $\sum_{v\in\mathcal{D}, v\neq y} v = -y$, it follows that $2f(x+y) \leq f(x) + f(x+2y)$ (which is the convexity of f in y), from which the second expression follows.

In fact, we just showed that the optimal actions are *transition monotone* [16], which means that every event occurring at an operator makes it more attractive for the other operators to choose the action that forces a state change. Conversely, if a transition occurs, it is less likely that it is still optimal for the same operator to take the action that makes the state change. We now use this monotonicity of the operators to prove the theorem. Consider some operator T_C with vector y. We have to show that

$$T_C f(x) + T_C f(x+v+w) \leq T_C f(x+v) + T_C f(x+w)$$

for all $v, w \in \mathcal{D}$. If the minimizing actions in $x+v$ and $x+w$ are equal, then it follows immediately by induction. For example, assume that the minimizing action is a state change; thus, $T_C f(x+v) + T_C f(x+w) = c + f(x+y+v) + c + f(x+y+w)$. Then,

$$\begin{aligned} T_C f(x) + T_C f(x+v+w) &\leq c + f(x+y) + c + f(x+y+v+w) \\ &\leq c + f(x+y+v) + c + f(x+y+w), \end{aligned}$$

the second inequality by $\mathrm{MM}(v,w)(x+y)$.

Assume that both minimizing actions are different and that $y \neq v, w$. Then, we can take $T_C f(x+v) + T_C f(x+w) = f(x+v) + c + f(x+y+w)$, without loss of generality. However, then,

$$\begin{aligned} T_C f(x) + T_C f(x+v+w) &\leq f(x) + c + f(x+y+v+w) \\ &\leq f(x+v) + c + f(x+y+w), \end{aligned}$$

the second inequality by Lemma 3.4.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

Finally, assume that $y = v$. By the monotonicity of the optimal action, we know that if the actions are different, then it must be such that $T_C f(x+v) + T_C f(x+w) = f(x+v) + c + f(x+y+w) = f(x+v) + c + f(x+v+w)$. It follows that

$$\begin{aligned} T_C f(x) + T_C f(x+v+w) &\le c + f(x+y) + f(x+v+w) \\ &= T_C f(x+v) + T_C f(x+w). \end{aligned}$$

The proof for T_{costs} and T_{env} is trivial. ■

Theorem 4.1 shows us that models consisting of the controlled operators corresponding to a tandem system (plus additional uncontrolled arrivals at any of the queues) has a mm value function. By Lemma 3.6, other properties such as convexity in all of the state components follow. Uncontrolled operators, however, cannot be propagated through mm alone. To see this, consider $T_{D(i)}$ and $f \in \mathcal{F}(\text{MM})$. As $\mathcal{F}(\text{MM}) \subset \mathcal{F}(\text{CC}(i))$, let us see what we need to have $T_{D(i)} \in \mathcal{F}(\text{CC}(i))$. For $x_i = 0$, we need $I(i)$, indicating also that mm is not enough for the study of uncontrolled operators. The first step in this direction is the following theorem.

THEOREM 4.2: *For all $f \in \mathcal{F}(F, I)$, if the costs of operating $c \ge 0$ for all operators $T_{CTS(1)}, \ldots, T_{CTS(m-1)}, T_{CD(m)}$, then*

$$T_{A(1)} f, \ldots, T_{A(m)} f, T_{CA(1)} f, \ldots, T_{CA(m)} f, T_{CTS(1)} f, \ldots, T_{CTS(m-1)} f, T_{CD(m)} f \in \mathcal{F}(F, I);$$

for all $f_1, \ldots, f_l \in \mathcal{F}(F, I)$, if $C \in \mathcal{F}(F, I)$, then

$$T_{\text{costs}} f, T_{\text{env}}(f_1, \ldots, f_l) \in \mathcal{F}(F, I).$$

PROOF: We consider, for each operator, the inequalities that constitute $\mathcal{F}(F, I)$, one by one. By Lemma 3.8, $\mathcal{F}(F(1), \ldots, F(m-1), I(m)) = \mathcal{F}(F, I)$; thus, we only have to consider $F(1), \ldots, F(m-1), I(m)$.

For $T_{A(1)}, \ldots, T_{A(m)}$, the result follows directly by induction. Consider $T_{CA(i)}$. Consider $F(j)$. We have to show $T_{CA(i)} f(x + e_{j+1}) \le T_{CA(i)} f(x + e_j)$; this follows by taking, on the left-hand side, the minimizing action of the right-hand side; for example, if it is optimal to accept a customer in $x + e_j$, then

$$T_{CA(i)} f(x + e_{j+1}) \le f(x + e_i + e_{j+1}) \le f(x + e_i + e_j) = T_{CA(i)} f(x + e_j).$$

The proof for $I(m)$ is equivalent. Consider $T_{CTS(i)}$. The proof is similar to that of $T_{CA(i)}$, except for $F(i)$, if $x_i = 0$ and the optimal action in $x + e_i$ is serving. Then,

$$T_{CTS(i)} f(x + e_{i+1}) = f(x + e_{i+1}) \le c + f(x + e_{i+1}) = T_{CA(i)} f(x + e_i).$$

The same line of reasoning works for $I(m)$ and $T_{CD(m)}$. The proof for T_{costs} and T_{env} is again trivial. ■

For $f \in \mathcal{F}(F, I)$, we have that $f(x - e_i + e_{i+1}) \le f(x)$ if $x_i > 0$ and $f(x - e_m) \le f(x)$ if $x_m > 0$; thus, if $c = 0$ for the operators $T_{CTS(1)}, \ldots, T_{CTS(m-1)}, T_{CD(m)}$, then it is always optimal to serve and, thus, the operators are equal to their uncontrolled equivalent. This leads to the following corollary.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

COROLLARY 4.3: *For all* $f \in \mathcal{F}(\mathrm{MM}, F, I)$,

$$T_{TS(1)} f, \ldots, T_{TS(m-1)} f, T_{D(m)} f \in \mathcal{F}(\mathrm{MM}, F, I).$$

Remark 4.4: The necessity to include $F(i)$ if the server at queue i is uncontrolled was already recognized in [16, Sec. 4]. However, to propagate $F(i)$, $F(i+1)$, $\ldots, F(m-1)$ and $I(m)$ are also necessary. In fact, Theorem 4.2 can be generalized slightly by assuming only $F(i), \ldots, F(m-1)$ and $I(m)$; then, the service operators can be uncontrolled from queue i onward.

We next move to the operators that model multiple servers. We first argue why the set of mm functions does not propagate. By Lemma 3.5, $\mathrm{SuperC}(i,j)$ holds if f is mm. Let us apply $T_{MD(i)}$ with $S = 2$ in state x with $x_i = 0$. Thus, we need to show

$$T_{MD(i)} f(x + e_i) + T_{MD(i)} f(x + e_i + e_j) \leq T_{MD(i)} f(x + 2e_i) + T_{MD(i)} f(x + e_j),$$

which is equal to

$$f(x) + f(x + e_i) + f(x + e_j) + f(x + e_i + e_j) \leq 2f(x + e_i) + 2f(x + e_j),$$

which gives, after cancellation of terms, *submodularity*—the inverse of supermodularity. Because supermodularity is a consequence of mm, this line of reasoning does not work. The solution is to relax the conditions on the value function and to show that for this larger class of functions, the operators related to multiple servers do propagate.

THEOREM 4.5: *For all* $f \in \mathcal{F}(\mathrm{DC}, F, I)$,

$$T_{A(1)} f, \ldots, T_{A(m)} f, T_{MTS(1)} f, \ldots, T_{MTS(m-1)} f, T_{MD(m)} f \in \mathcal{F}(\mathrm{DC}, F, I),$$

and if $C \in \mathcal{F}(\mathrm{DC}, F, I)$ *and* $f_1, \ldots, f_l \in \mathcal{F}(\mathrm{DC}, F, I)$,

$$T_{\mathrm{costs}} f, T_{\mathrm{env}}(f_1, \ldots, f_l) \in \mathcal{F}(\mathrm{DC}, F, I).$$

PROOF: As usual, the proof for the operators $T_{A(1)}, \ldots, T_{A(m)}$, T_{costs}, and T_{env} is simple. Let us concentrate on the servers. Of course, because of Lemmas 3.2 and 3.8, we only have to consider $\mathrm{CC}(i)$ for all i, $\mathrm{Super}(i,j)$ for all $i \neq j$, $F(1), \ldots, F(m-1)$, and $I(m)$. The proof for $F(i)$ and $I(m)$ is simple. Consider $T_{MTS(k)}$. Propagating $\mathrm{CC}(i)$ for $i \neq k$ is trivial; consider $i = k$. We consider the cases $S \leq x_i$, $S = x_i + 1$, and $S \geq x_i + 2$.

- If $S \leq x_i$, then all S servers are busy in x, $x + e_i$, and $x + 2e_i$, and, thus,

$$\begin{aligned} 2T_{MTS(i)} f(x + e_i) &= 2Sf(x + e_{i+1}) \\ &\leq Sf(x - e_i + e_{i+1}) + Sf(x + e_i + e_{i+1}) \\ &= T_{MTS(i)} f(x) + T_{MTS(i)} f(x + 2e_i), \end{aligned}$$

  where the inequality holds by applying $\mathrm{CC}(i)$ in $x - e_i + e_{i+1}$ for S times, which we write as $S \times \mathrm{CC}(i)(x - e_i + e_{i+1})$.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

- If $S = x_i + 1$, then we get

$$\begin{aligned} 2T_{MTS(i)} f(x + e_i) &= 2Sf(x + e_{i+1}) \\ &\le (S-1)f(x - e_i + e_{i+1}) + f(x) + Sf(x + e_i + e_{i+1}) \\ &= T_{MTS(i)} f(x) + T_{MTS(i)} f(x + 2e_i). \end{aligned}$$

  The inequality is equal to $(S-1) \times \mathrm{CC}(i)(x - e_i + e_{i+1}) + \mathrm{Super}(i, i+1)(x) + F(i)(x)$. Note that if $x_i = 0$ and, thus, $S = 1$, then the coefficient of $\mathrm{CC}(i)(x - e_i + e_{i+1})$ is zero, avoiding negative queue lengths.
- If $S \ge x_i + 2$, then

$$\begin{aligned} 2T_{MTS(i)} f(x + e_i) &= 2(x_i + 1)f(x + e_{i+1}) + 2(S - x_i - 1)f(x + e_i) \\ &\le x_i f(x - e_i + e_{i+1}) + (S - x_i)f(x) \\ &\quad + (x_i + 2)f(x + e_i + e_{i+1}) + (S - x_i - 2)f(x + 2e_i) \\ &= T_{MTS(i)} f(x) + T_{MTS(i)} f(x + 2e_i), \end{aligned}$$

  which follows from $x_i \times \mathrm{CC}(i)(x - e_i + e_{i+1}) + (S - x_i - 2) \times \mathrm{CC}(i)(x) + 2\,\mathrm{Super}(i, i+1)(x)$. Again, the coefficients are such that $\mathrm{CC}(i)(x - e_i + e_{i+1})$ is only used if $x_i > 0$.

We now propagate $T_{MTS(k)}$ for $\mathrm{Super}(i, j)$. Again, $k \ne i, j$ is trivial; assume that $k = i$. First assume that $j \ne i + 1$. We consider the cases $S \le x_i$ and $S \ge x_i + 1$.

- If $S \le x_i$, then all S servers are busy in all states. This gives, by $S \times \mathrm{Super}(i, j)(x - e_i + e_{i+1})$,

$$\begin{aligned} T_{MTS(i)} f(x + e_i) + T_{MTS(i)} f(x + e_j) &= Sf(x + e_{i+1}) + Sf(x - e_i + e_{i+1} + e_j) \\ &\le Sf(x - e_i + e_{i+1}) + Sf(x + e_{i+1} + e_j) \\ &= T_{MTS(i)} f(x) + T_{MTS(i)} f(x + e_i + e_j). \end{aligned}$$

- If $S \ge x_i + 1$, then

$$\begin{aligned} &T_{MTS(i)} f(x + e_i) + T_{MTS(i)} f(x + e_j) \\ &\quad = (x_i + 1)f(x + e_{i+1}) + (S - x_i - 1)f(x + e_i) \\ &\qquad + x_i f(x - e_i + e_{i+1} + e_j) + (S - x_i)f(x + e_j) \\ &\quad \le x_i f(x - e_i + e_{i+1}) + (S - x_i)f(x) + (x_i + 1)f(x + e_{i+1} + e_j) \\ &\qquad + (S - x_i - 1)f(x + e_i + e_j) \\ &\quad = T_{MTS(i)} f(x) + T_{MTS(i)} f(x + e_i + e_j). \end{aligned}$$

  The inequality holds because of $x_i \times \mathrm{Super}(i, j)(x - e_i + e_{i+1}) + (S - x_i - 1) \times \mathrm{Super}(i, j)(x) + \mathrm{Super}(i + 1, j)(x)$.

Now, assume that $k = i$ and $j = i + 1$. We consider the cases $S \le x_i$ and $S \ge x_i + 1$.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

- If $S \leq x_i$, then all S servers are busy in all states. This gives, by $S \times \text{Super}(i, i+1)(x - e_i + e_{i+1})$,

$$\begin{aligned} &T_{MTS(i)} f(x + e_i) + T_{MTS(i)} f(x + e_j) \\ &\quad = Sf(x + e_{i+1}) + Sf(x - e_i + 2e_{i+1}) \\ &\quad \leq Sf(x - e_i + e_{i+1}) + Sf(x + 2e_{i+1}) \\ &\quad = T_{MTS(i)} f(x) + T_{MTS(i)} f(x + e_i + e_j). \end{aligned}$$

- If $S \geq x_i + 1$, then

$$\begin{aligned} &T_{MTS(i)} f(x + e_i) + T_{MTS(i)} f(x + e_j) \\ &\quad = (x_i + 1) f(x + e_{i+1}) + (S - x_i - 1) f(x + e_i) + x_i f(x - e_i + 2e_{i+1}) \\ &\qquad + (S - x_i) f(x + e_{i+1}) \\ &\quad \leq x_i f(x - e_i + e_{i+1}) + (S - x_i) f(x) + (x_i + 1) f(x + 2e_{i+1}) \\ &\qquad + (S - x_i - 1) f(x + e_i + e_{i+1}) \\ &\quad = T_{MTS(i)} f(x) + T_{MTS(i)} f(x + e_i + e_j). \end{aligned}$$

  The inequality holds because of $x_i \times \text{Super}(i, i+1)(x - e_i + e_{i+1}) + (S - x_i - 1) \times \text{Super}(i, i+1)(x) + \text{CC}(i+1)(x)$.

Now, consider $T_{MD(m)}$—first for $CC(m)$. We consider the cases $S \leq x_m$, $S = x_m + 1$, and $S \geq x_m + 2$.

- If $S \leq x_m$, then all S servers are busy in x, $x + e_m$, and $x + 2e_m$ and, thus,

$$\begin{aligned} 2T_{MTS(m)} f(x + e_m) &= 2Sf(x) \\ &\leq Sf(x - e_m) + Sf(x + e_m) \\ &= T_{MTS(m)} f(x) + T_{MTS(m)} f(x + 2e_m); \end{aligned}$$

  the inequality equals $S \times \text{CC}(m)(x - e_m)$.
- If $S = x_m + 1$, then we get

$$\begin{aligned} 2T_{MTS(m)} f(x + e_m) &= 2Sf(x) \\ &\leq (S - 1) f(x - e_m) + f(x) + Sf(x + e_m) \\ &= T_{MTS(m)} f(x) + T_{MTS(m)} f(x + 2e_m); \end{aligned}$$

  the inequality is equal to $(S - 1) \times \text{CC}(m)(x - e_m) + I(m)(x)$.
- If $S \geq x_m + 2$, then

$$\begin{aligned} 2T_{MTS(m)} f(x + e_m) &= 2(x_m + 1) f(x) + 2(S - x_m - 1) f(x + e_m) \\ &\leq x_m f(x - e_m) + (S - x_m) f(x) + (x_m + 2) f(x + e_m) \\ &\qquad + (S - x_m - 2) f(x + 2e_m) \\ &= T_{MTS(m)} f(x) + T_{MTS(m)} f(x + 2e_m), \end{aligned}$$

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

where the inequality equals $x_m \times \mathrm{CC}(m)(x - e_m) + (S - x_m - 2) \times \mathrm{CC}(m)(x)$. Proving $\mathrm{CC}(i)$ for $i \neq m$ is simple.

Finally, we prove $T_{MD(m)}$ for $\mathrm{Super}(i, j)$. Assume that $i = m$, the other case being simple. It is readily seen that only $\mathrm{Super}(m, j)$ suffices to prove the inequalities. ■

We summarize the results of this section in the following corollary.

COROLLARY 4.6: *For V_n constructed from $T_{A(1)}, \ldots, T_{A(m)}$, $T_{CA(1)}$, $T_{CTS(1)}, \ldots, T_{CTS(m-1)}$, $T_{CD(m)}$, T_{env}, and T_{costs} with $C \in \mathcal{F}(\mathrm{MM})$ and if $V_0 \in \mathcal{F}(\mathrm{MM})$, then $V_n \in \mathcal{F}(\mathrm{MM})$ for all $n > 0$.*

For V_n constructed from $T_{A(1)}, \ldots, T_{A(m)}$, $T_{CA(1)}$, $T_{CTS(1)}, \ldots, T_{CTS(m-1)}$ with $c > 0$, $T_{CD(m)}$ with $c > 0$, $T_{TS(1)}, \ldots, T_{TS(m-1)}$, $T_{D(m)}$, T_{env}, and T_{costs} with $C \in \mathcal{F}(\mathrm{MM}, F, I)$, and if $V_0 \in \mathcal{F}(\mathrm{MM}, F, I)$, then $V_n \in \mathcal{F}(\mathrm{MM}, F, I)$ for all $n > 0$.

For V_n constructed from $T_{A(1)}, \ldots, T_{A(m)}$, $T_{MTS(1)}, \ldots, T_{MTS(m-1)}$, $T_{MD(m)}$, T_{env}, and T_{costs} with $C \in \mathcal{F}(\mathrm{DC}, F, I)$ and if $V_0 \in \mathcal{F}(\mathrm{DC}, F, I)$, then $V_n \in \mathcal{F}(\mathrm{DC}, F, I)$ for all $n > 0$.

Up to now, we have not really discussed the crucial role played by the direct costs C of T_{costs}. As we can take any function C satisfying the right set of inequalities, there are many possibilities. An obvious choice is $C(x) = \sum_i c_i x_i$, with $c_1 \geq \cdots \geq c_m$ if $C \in \mathcal{F}(F)$. If $C \in \mathcal{F}(\mathrm{MM})$, we can also take $C(x) = \sum_i (x_i - B_i)^+ P_i$, introducing a penalty P_i for each customer that exceeds the buffer level B_i in queue i.

5. APPLICATIONS

We provide some applications of the comparison of models with different input processes. As a simple first application, consider two models, with value function V_n and V_n', that are different only in the fact that the first has an operator $T_{A(1)}$ and the latter has $\frac{1}{2}T_I + \frac{1}{2}T_{A(1)}T_{A(1)}$. Here, T_I is simply the identity operator [i.e., $T_I f(x) = f(x)$]. We did not discuss it yet, but it is clear that if $f \in \mathcal{F}$ for some set $\mathcal{F}$, then also $T_I f \in \mathcal{F}$. An operator that models convex combinations is easily built with T_{env}; thus, $\frac{1}{2}T_I + \frac{1}{2}T_{A(1)}T_{A(1)}$ falls within our framework. In words, under V_n, arrivals at queue 1 are twice more likely to occur than under V_n', but each arrival in V_n' is a batch of size 2. Thus, the resulting load is equal. We would like to show that $V_n(x) \leq V_n'(x)$. Propagating this for all operators of V_n other than $T_{A(1)}$ is trivial, as the same operators occur in V_n'. Thus, let us try to show

$$T_{A(1)} V_n(x) \leq (\tfrac{1}{2}T_I + \tfrac{1}{2}T_{A(1)}T_{A(1)}) V_n'(x). \tag{3}$$

This is equivalent to

$$V_n(x + e_1) \leq \tfrac{1}{2}V_n'(x) + \tfrac{1}{2}V_n'(x + 2e_1).$$

Now, assume that V_n is convex in x_1. Then,

$$V_n(x + e_1) \leq \tfrac{1}{2}V_n(x) + \tfrac{1}{2}V_n(x + 2e_1) \leq \tfrac{1}{2}V_n'(x) + \tfrac{1}{2}V_n'(x + 2e_1)$$

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

and, thus, (3) holds. Therefore, if the value function is convex in the first component, the model with batch arrivals has greater costs. To check that the value function satisfies CC(1), we use Corollary 4.6, noting that $\mathcal{F}(\mathrm{DC}) \subset \mathcal{F}(\mathrm{CC}(1))$. This also imposes conditions on the performance measure; for example, if all service operators are controlled, then we have the condition $C \in \mathcal{F}(\mathrm{MM})$, which allows for decreasing direct costs. On the other hand, if the operators are uncontrolled, then we need $C \in \mathcal{F}(\mathrm{MM}, F, I)$ or $C \in \mathcal{F}(\mathrm{DC}, F, I)$ and, thus, the direct costs must be increasing in all components.

This example is a nice illustration of how convexity in the first state component can be used to compare systems with different arrival processes. More complex systems can be handled as well, as we show next. We consider the model of [9], in which a single queue is considered with an arrival process consisting of parallel on–off processes. In [9], the number of on–off processes is varied (while keeping the total offered load equal), and it is shown that systems with many on–off sources perform better for direct costs for which the value function is convex. See [9, Lemma 3.4]. We generalize this Lemma 3.4 to tandems of queues.

To do so, we first show how to model N on–off sources (in fact, the whole value function) with the environment operator. Consider first the states. The set of all possible states of the on–off sources is $\{0,\dots,K\}$, representing the number of on–off sources that is on. If necessary, we could model an additional environment variable (note that a multidimensional environment can always be written as a one-dimensional environment, because the number of states remains countable), but we will not do so. By $x \in \mathbb{N}_0^m$, we denote the state of the queues. As in [9], p is the probability that an on–off source goes on and q is the probability that an on–off source goes off. When a source is on, customers arrive at a rate λ. Furthermore, assume that the service rates of queue 1 up to m are equal to $\mu_1,\dots,\mu_m$ and assume that $(p+q+\lambda)K + \sum_i \mu_i = 1$. Then,

$$\begin{aligned} T_{\mathrm{env}}^K(f_I, f_A, f_1, \dots, f_m)(k,x) = {} & (K-k)pf_I(k+1,x) + kqf_I(k-1,x) \\ & + (kp + (K-k)(q+\lambda))f_I(k,x) + k\lambda f_A(k,x) \\ & + \sum_i \mu_i f_i(k,x). \end{aligned}$$

Denote with $V_n^K(k,x)$ the value function at stage n of the system with K sources. (With respect to the notation in [9], we dropped the environment variable x and replaced N, n, and k by K, k, and n, respectively.) Using the definition of T_{env}^K and writing T_i as a generic way to indicate the service operator at queue i, we have the following recursive expression for $V_n^K(k,x)$:

$$V_{n+1}^K(k,x) = T_{\mathrm{costs}}[T_{\mathrm{env}}^K(V_n, T_{A(1)}V_n, T_iV_n, \dots, T_mV_n)](k,x).$$

Having modeled the tandem system with an on–off source input, we are ready to compare V_n^K with V_n^1. The equivalent of equation (20) in [9] for tandem systems, which we have to show here, is

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

$$\sum_{k=0}^{K}\binom{K}{k}p^k q^{K-k}((K-k)V_n^K(k,x)+kV_n^K(k,x+y))$$
$$\leq K(p+q)^{K-1}qV_n^1(0,x)+K(p+q)^{K-1}pV_n^1(1,x+y) \quad \textbf{(4)}$$

for all $x, y \geq 0$.

The comparison result follows by taking $y = 0$ and observing that the coefficients are proportional to the stationary probabilities of the on–off sources. Thus the comparison holds for an initially stationary arrival process. To prove (4), we first have to assume that $V_n^K \in \mathcal{F}(\mathrm{DC})$. We discuss those parts of the proof of (4) that are different from the proof given in [9] for a single dimension. To prove the terms related to arriving customers, we need

$$(K-k)^2V_n^K(k,x)+(K-k)kV_n^K(k,x+e_1)+k(K-k)V_n^K(k,x+y)$$
$$+k^2V_n^K(k,x+y+e_1)$$
$$\leq (K-k)KV_n^K(k,x)+kKV_n^K(k,x+y+e_1).$$

This holds (after cancellation of terms) by directional convexity of the value function.

To propagate the tandem servers, we have to assume that the operators are uncontrolled, thus $T_{MTS(i)}$ for $i < m$ and $T_{MD(m)}$. Otherwise, it might occur that the minimizing action on the right-hand side of (4) is different. In conclusion, Lemma 3.4 of [9] and, with that, the related Theorems 3.5 and 3.6 of [9] are generalized to tandem systems with uncontrolled operators and cost functions that belong to $\mathcal{F}(\mathrm{DC}, F, I)$.

Directional convexity is also useful for certain minimization problems involving multiple systems. Here, we consider customers that have to be handled by one of a group of parallel tandem systems. The arrival stream has to be split up in a Bernoulli fashion. Convexity of the objective function allows that the global minimum can be found using a simple minimization procedure such as steepest descent. To prove this convexity, we have to show that the performance measure for a single tandem system is convex in the rate assigned to that system. We will prove this, again using directional convexity in the state. Define the operator $T_{A(i)}^{\lambda}$, for $\lambda \in [0,1]$, by $T_{A(i)}^{\lambda} f(x) = \lambda f(x+e_i) + \bar{\lambda} f(x)$ (with $\bar{\lambda} = 1-\lambda$) and let $V_n^{\lambda}(x)$ be the value function of a tandem system with uncontrolled operators as defined in Section 2 and by $T_{A(i)}^{\lambda}$. We would like to show convexity in λ; that is,

$$V_n^{\alpha\lambda_1+\bar{\alpha}\lambda_2}(x) \leq \alpha V_n^{\lambda_1}(x) + \bar{\alpha}V_n^{\lambda_2}(x). \quad \textbf{(5)}$$

If we assume $\lambda_1 \leq \lambda_2$, then for each of these value functions, if $T_{A(i)}^{\cdot}$ occurs, then the arrivals occur with a probability in $[\lambda_1, \lambda_2]$. Thus, we can rewrite $T_{A(i)}^{\beta}$, $\beta \in [\lambda_1, \lambda_2]$, as $T_{A(i)}^{\beta} = \lambda_1 T_{A(i)} + (\lambda_2 - \lambda_1)T_{A(i)}^{\gamma} + \bar{\lambda}_2 T_I$, with $\gamma = (\beta - \lambda_1)/(\lambda_2 - \lambda_1)$. Note that if $\beta \in [\lambda_1, \lambda_2]$, then $\gamma \in [0,1]$. Thus, proving (5) for all α is, given that the standard arrival operator propagates (which is the case for all sets of functions that we considered), equivalent to proving

$$V_n^{\alpha}(x) \leq \alpha V_n^1(x) + \bar{\alpha}V_n^0(x)$$

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

for all α. This inequality follows from

$$\alpha V_n^\alpha(x+y) + \bar{\alpha} V_n^\alpha(x) \leq \alpha V_n^1(x+y) + \bar{\alpha} V_n^0(x),$$

which we will prove for all $y \geq 0$. Note the similarity with (4). Again, the only operator that is not trivial to propagate is $T_{A(i)}^{\cdot}$; it results in

$$\begin{aligned}\alpha^2 V_n^\alpha(x+y+e_i) + \alpha\bar{\alpha} V_n^\alpha(x+y) + \bar{\alpha}\alpha V_n^\alpha(x+e_i) + \bar{\alpha}^2 V_n^\alpha(x+y)\\ \leq \alpha V_n^1(x+y+e_i) + \bar{\alpha} V_n^0(x).\end{aligned}$$

It suffices to show that the left-hand side is smaller than $\alpha V_n^\alpha(x+y+e_i) + \bar{\alpha} V_n^\alpha(x)$, which is equivalent to

$$V_n^\alpha(x+y) + V_n^\alpha(x+e_i) \leq V_n^\alpha(x+y+e_i) + V_n^\alpha(x),$$

which holds by the directional convexity of V_n^α. Thus, for the same systems as the last example with on–off sources, we have shown that the performance is convex in the arrival rate. This generalizes part (i) of Theorem 2.7 of [11] to multistage systems.

Acknowledgments

I thank Ludolf Meester for some interesting discussions and additional insights while we were (in vain) trying to bridge the gap between our methods. I also would like to thank a referee of an earlier version of this paper for his detailed comments.

References

1. Altman, E., Gaujal, B., & Hordijk, A. (2000). Multimodularity, convexity and optimization properties. *Mathematics of Operations Research* 25: 324–347.
2. Altman, E. Jiménez, T., & Koole, G.M. (2001). On the comparison of queueing systems with their fluid limits. *Probability in the Engineering and Informational Sciences* 15: 165–178.
3. Altman, E. & Koole, G.M. (1998). On submodular value functions and complex dynamic programming. *Stochastic Models* 14: 1051–1072.
4. Asmussen, S. & Koole, G.M. (1993). Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability* 30: 365–372.
5. Chang, C.S., Chao, X., Pinedo, M., & Weber, R.R. (1992). On the optimality of LEPT and $c\mu$-rules for machines in parallel. *Journal of Applied Probability* 29: 667–681.
6. Hajek, B. (1985). Extremal splitting of point processes. *Mathematics of Operations Research* 10: 543–556.
7. Hordijk, A. & Koole, G.M. (1992). On the assignment of customers to parallel queues. *Probability in the Engineering and Informational Sciences* 6: 495–511.
8. Koole, G.M. (1998). Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems* 30: 323–339.
9. Koole, G.M. & Liu, Z. Stochastic bounds for queueing systems with multiple on–off sources. *Probability in the Engineering and Informational Sciences* 12: 25–48.
10. Lippman, S.A. (1975). Applying a new device in the optimization of exponential queueing systems. *Operations Research* 23: 687–710.
11. Liyanage, L. & Shanthikumar, J.G.(1992). Second-order properties of single-stage queueing systems. In U.N. Bhat and I.S. Basawa (eds.), *Queueing and related models*. Oxford: Oxford University Press, pp. 129–160.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press

12. Meester, L.E. (1990). Contributions to the theory and applications of stochastic convexity. Ph.D. thesis, University of California at Berkeley.
13. Puterman, M.L. (1994). *Markov decision processes*. New York: Wiley.
14. Shaked, M. & Shanthikumar, J.G. (1990). Parametric stochastic convexity and concavity of stochastic processes. *Annals of the Institute of Statistical Mathematics* 42: 509–531.
15. Stidham, S., Jr. & Weber, R.R. (1993). A survey of Markov decision models for control of networks of queues. *Queueing Systems* 13: 291–314.
16. Weber, R.R. & Stidham, S., Jr. (1987). Optimal control of service rates in networks of queues. *Advances in Applied Probability* 19: 202–218.

https://doi.org/10.1017/S0269964804181023 Published online by Cambridge University Press