J. R. Statist. Soc. B (1999)
61, *Part* 4, *pp.* 747–791

The achievable region approach to the optimal control of stochastic systems

Marcus Dacre and Kevin Glazebrook

University of Newcastle, UK

and José Niño-Mora

Universitat Pompeu Fabra, Barcelona, Spain

[*Read before* The Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday, February 10th, 1999*, Professor P. J. Green *in the Chair*]

Summary. The achievable region approach seeks solutions to stochastic optimization problems by characterizing the space of all possible performances (the achievable region) of the system of interest and optimizing the overall system-wide performance objective over this space. This is radically different from conventional formulations based on dynamic programming. The approach is explained with reference to a simple two-class queuing system. Powerful new methodologies due to the authors and co-workers are deployed to analyse a general multiclass queuing system with parallel servers and then to develop an approach to optimal load distribution across a network of interconnected stations. Finally, the approach is used for the first time to analyse a class of intensity control problems.

Keywords: Achievable region; Gittins index; Linear programming; Load balancing; Multiclass queuing systems; Performance space; Stochastic optimization; Threshold policy

1. Introduction

The last decade has seen a substantial research focus on the modelling, analysis and optimization of complex stochastic service systems, motivated in large measure by applications in areas such as computer and telecommunication networks. Optimization issues which broadly focus on making the best use of limited resources are recognized as of increasing importance. However, stochastic optimization in the context of systems of any complexity is technically very difficult.

For the most part, the optimal dynamic control of queuing and other stochastic systems has been approached via dynamic programming (DP) formulations. Within such formulations, a variety of special arguments (of which the simplest and most effective have been interchange arguments) have been adduced to obtain structural results concerning optimal controls. A good summary of how things stood in the mid- to late 1980s can be found in chapters 8 and 9 of Walrand (1988). It would not be unfair to claim that a consensus view of this enterprise is that there was relatively little to show for a large amount of effort and that a pressing need existed for new approaches. One notable success, though, was the elucidation by Gittins (1979, 1989) of index-based solutions to a variety of optimization problems

Address for correspondence: Kevin Glazebrook, Department of Statistics, University of Newcastle, Newcastle upon Tyne, NE1 7RU, UK.
E-mail: kevin.glazebrook@newcastle.ac.uk

 1369–7412/99/61747

concerned with the sequential allocation of effort among a collection of competing projects (or *bandits* as they are sometimes called). In Gittins's solutions, each project has a priority index which depends on its current state (and, possibly, its past history) and effort is optimally allocated at each decision epoch to whichever project has the highest current index value. The indices concerned are now known as *Gittins indices* and the associated procedure for the allocation of effort is a *Gittins index* (or, simply, an index) policy. Until fairly recently, most subsequent developments of index theory by, for example, Glazebrook (1982), Weber (1992), Weiss (1988) and Whittle (1980) took place within DP formulations (broadly defined). *Inter alia*, index policies have been shown to be optimal for a range of simple multiclass queuing systems. In these, an index is typically attached to each customer class. The resulting index policy will select customers for service on the basis of (maximal index) class membership only.

Many of the most important recent developments in the control, for example, of multiclass queuing networks have sought to optimize some associated (limiting) process, whether a diffusion process (Brownian system model) in heavy traffic (see, for example, Harrison and Nguyen (1993) and Harrison and Wein (1989)) or a fluid model (see, for example, Atkins and Chen (1995) and Maglaras (1997)). These are powerful methodologies and have rightly been very influential. However, since the main focus of optimization is an approximating or limiting process there can be formidable challenges in the subsequent derivation of controls for the queuing system of original interest and in the evaluation of such controls. See Harrison (1996) and Maglaras (1997).

The paper concerns a different approach — namely the so-called achievable region or mathematical programming approach. It is possible that this could ultimately turn out to be more limited in its range of applications than those cited above (although the current pace of development throughout the field makes a final judgment impossible). However, it does have the considerable advantage of staying in firm contact with the original stochastic system of interest throughout. Hence, when analyses via this methodology are available, they typically make clear and strong statements about the control policies identified.

The achievable region approach seeks solutions to stochastic optimization problems by

(a) characterizing the space of all possible performances (the achievable region) of the stochastic system and
(b) optimizing the overall system-wide performance objective over this space.

The performance space in (a) is often a polyhedron of special structure, yielding in (b) a mathematical program (usually a linear program (LP)) for which efficient algorithms exist. In some respects, the approach can be thought of as having its roots in well-established LP approaches to stochastic optimization; see, for example, Manne (1960). However, rather than use standard LP formulations in the variable space of state–action frequencies (which is typically huge or infinite) we aim to develop analyses in some *projected* space (hopefully, of much reduced dimensionality) of *natural* performance variables. The earliest contributions in this vein were due to Gelenbe and Mitrani (1980), followed by Federgruen and Groenevelt (1988). Contributions by Shanthikumar and Yao (1992) and Bertsimas and Niño-Mora (1996) took the approach decisively further forward, the latter giving an account of Gittins indices from this perspective. In an important related contribution, Whittle (1988) considered the intractable restless bandit problem, in which the competing projects models of Gittins (1979) were generalized to allow for changes of state to occur in projects even when not in receipt of effort. Whittle's analysis via an LP relaxation of the problem produced an index characterized as a Lagrange multiplier associated with a conservation constraint. Weber and

Weiss (1990) developed these ideas further and established a form of asymptotic optimality of the policy based on Whittle's index under stated conditions.

Our goal is, firstly, to bring the achievable region approach to the attention of a wider audience than it has enjoyed hitherto. For this, many of the ideas alluded to in the previous paragraph are presented in Section 2 in a way which we trust will be widely accessible. In addition a range of powerful new methodologies with which the authors and co-workers have been associated are described and illustrated by the discussions in Sections 3–5 of a range of important stochastic optimization problems. This material is new and should convey something of the power and scope of the achievable region approach. Given a familiarity with the content of Section 2, the later sections are self-contained with Section 3 the most demanding technically. Section 3 discusses the status of index policies for a general multiclass queuing system with servers working in parallel. We consider in Section 4 an approach to distributing the workload across a network of interconnected stations when each station is assumed to schedule its own offered load optimally. The problem of controlling input and output rates for a simple queuing system is discussed in Section 5. The paper concludes in Section 6 with proposals for future work.

2. The achievable region approach

For definiteness, we shall develop the core ideas underlying the achievable region approach in the context of multiclass queuing systems. Such systems have frequently been proposed as models for computer and communication systems which typically are requested to handle several traffic types simultaneously. Let $E = \{1, 2, \ldots, N\}$ denote a set of *customer classes*. *Customers* in the system require service which is provided by a collection of *servers*. A *control* u is a rule for determining how servers should be assigned to waiting customers. The set of *admissible controls* is denoted $\mathcal{U}$. Although admissibility will be defined in context, it will invariably be required that controls should be *non-anticipative* (decisions are made on the basis of the history of the process only) and *non-idling* (servers should never be idle when there is work to be done). With each control u is associated a *system performance vector* $\mathbf{x}^u = (x_1^u, x_2^u, \ldots, x_N^u)$ with x_i^u denoting the *class i performance*, $i \in E$. Throughout the paper, x_i^u will be the expectation of some quantity related to class i. A standard choice for x_i^u, denoted $E_u(N_i)$, is the long-term average number of class i customers in the system under control u. The *performance space* is the set of all possible performances, denoted $X = \{\mathbf{x}^u, u \in \mathcal{U}\}$. There is a cost $c(\mathbf{x}^u)$ associated with operating the system under control u which depends on the control only through its associated performance. The stochastic optimization problem of interest is expressed as

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} \{c(\mathbf{x}^u)\}. \tag{1}$$

The prime goal is the identification of a control u^{OPT} attaining the infimum in equation (1). If X is known, an alternative computation of Z^{OPT} is via the minimization

$$Z^{\mathrm{OPT}} = \inf_{\mathbf{x} \in X} \{c(\mathbf{x})\}. \tag{2}$$

In all the cases that we shall consider we shall have $c(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x}$ for some cost vector $\mathbf{c}$ and X a convex polyhedron, yielding in equation (2) an LP. Solving equation (2) will yield $\mathbf{x}^{\mathrm{OPT}}$, the optimizing performance. The question then arises of whether a control u^{OPT} can be found which realizes $\mathbf{x}^{\mathrm{OPT}}$.

The *achievable region approach* seeks solutions to stochastic optimization problems as in equation (1) by

(a) the identification of the performance space X,
(b) solving a mathematical programming problem as in equation (2) with feasible space X and
(c) the identification of controls yielding the optimum performance.

To give the reader some idea of how the approach might proceed, we outline the case of a two-class $M/M/1$ queuing system, first analysed in this manner by Coffman and Mitrani (1980).

Customers of class k arrive at a single server according to independent Poisson streams of rate λ_k with service requirements (independent of each other and of the arrival streams) which are exponentially distributed with mean μ_k^{-1}, $k = 1, 2$. The rate at which work arrives in the system is $\lambda_1/\mu_1 + \lambda_2/\mu_2$ which is assumed to be less than 1 (the available service rate) to guarantee stability, i.e. the time-average number of customers in the system is finite. Controls for the system must be non-anticipative and non-idling and priorities between customer classes may be imposed pre-emptively (i.e. a customer whose requirements have not yet been fully met may be removed from service to make way for another customer of higher priority). The goal is to choose a control u to minimize a long-term holding cost rate, i.e.

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} \{c_1\, E_u(N_1) + c_2\, E_u(N_2)\}. \tag{3}$$

In expression (3) c_k is a cost rate, N_k is the number of class k customers in the system and E_u denotes an expectation taken under the steady state distribution when control u is applied. The achievable region approach solves the stochastic optimization problem (3) by proceeding through the above steps (a)–(c) as follows.

2.1. *Identification of the performance space X*

Our first goal is to develop a suitable collection of equations and inequalities involving the steady state expectations $E_u(N_1)$ and $E_u(N_2)$. For this, we consider the work-in-system process whose value at time t under control u is denoted by $V_u(t)$. This quantity is equal to the sum of the remaining service times of all customers in the system at time t. All sample paths of $V_u(t)$ under non-idling controls u consist of upward jumps at customer arrival epochs (when additional service requirement enters the system) followed by a period during which this quantity is reduced at rate 1 as service is executed. Each such period terminates *either* with an empty system when $V_u(t)$ hits the value 0 *or* with an upward jump when the next customer arrives. A little reflection will enable the reader to see that the sample paths of $V_u(t)$ do not depend on the choice of non-idling u. It will certainly follow that in the steady state the expected work in the system is control invariant. In the case of our simple two-class system the constant concerned is easily identified and we can infer that

$$\frac{E_u(N_1)}{\mu_1} + \frac{E_u(N_2)}{\mu_2} = \frac{\rho_1\mu_1^{-1} + \rho_2\mu_2^{-1}}{1 - \rho_1 - \rho_2}, \qquad u \in \mathcal{U}. \tag{4}$$

That the expression on the left-hand side of equation (4) is the expected work-in-system in the steady state follows from the assumption of exponential service times. This implies that each class k customer in the system must have expected remaining service requirement equal to μ_k^{-1}, $k = 1, 2$. Note that the key quantity ρ_k in equation (4) is equal to λ_k/μ_k, $k = 1, 2$.

We develop these ideas further by considering $V_u^1(t)$ as the work-in-system at time t due to customers of class 1 only. We should not expect this new quantity to be control invariant. It will depend in general on how control u distributes service between the two classes. In fact, it is not difficult to see that, for each realization of the system, $V_u^1(t)$ will be minimized for each t by the control u which always gives class 1 customers priority in service over class 2. Hence the steady state expected class 1 work in the system is also minimized by such a priority policy, which we denote $1 \to 2$. This yields the inequality

$$\frac{E_u(N_1)}{\mu_1} \geqslant \frac{\rho_1 \mu_1^{-1}}{1-\rho_1}, \qquad u \in \mathcal{U}, \tag{5}$$

where the right-hand side of inequality (5) is the mean work-in-system of an $M/M/1$-system serving class 1 customers only. Repeating the argument from the perspective of class 2 we obtain

$$\frac{E_u(N_2)}{\mu_2} \geqslant \frac{\rho_2 \mu_2^{-1}}{1-\rho_2}, \qquad u \in \mathcal{U}, \tag{6}$$

with the right-hand side of inequality (6) attained when the system is controlled by the priority policy $2 \to 1$. Motivated by expressions (4)–(6), we take $x_k^u = E_u(N_k)/\mu_k$ as the class k performance associated with control u, $k = 1, 2$. From expressions (4)–(6), it immediately follows that performance space $X = \{(x_1^u, x_2^u), u \in \mathcal{U}\}$ is contained within the line segment P given by

$$P = \left\{(x_1, x_2);\ x_1 \geqslant \frac{\rho_1 \mu_1^{-1}}{1-\rho_1},\ x_2 \geqslant \frac{\rho_2 \mu_2^{-1}}{1-\rho_2},\ x_1 + x_2 = \frac{\rho_1 \mu_1^{-1} + \rho_2 \mu_2^{-1}}{1-\rho_1-\rho_2}\right\}. \tag{7}$$

See Fig. 1. To show that $P \subseteq X$, observe that the end points of P, labelled A and B, may be identified as the performances associated with the priority policies $1 \to 2$ and $2 \to 1$ respectively. This follows from the discussion around inequalities (5) and (6). Any point of P is a convex combination of A and B and hence is easily seen to be the performance of a suitable randomization of the policies $1 \to 2$ and $2 \to 1$. Hence all points of P are performances, as required. We conclude that $X = P$.

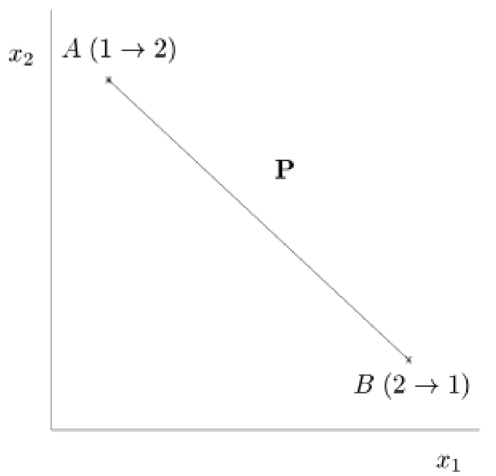

Fig. 1. Line segment P

2.2. Solution of a linear program with feasible space X

For reasons that will soon become clear, consider the LP

$$Z^{\mathrm{OPT}} = \inf_{\mathbf{x} \in P} (c_1 \mu_1 x_1 + c_2 \mu_2 x_2). \tag{8}$$

It is trivial to show that the minimum is attained at end point A when $c_1\mu_1 \geqslant c_2\mu_2$ and at end point B otherwise.

2.3. Solution of the stochastic optimization problem of interest

Our objective is to identify a control u^{OPT} to solve equation (3), rewritten as

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} (c_1 \mu_1 x_1^u + c_2 \mu_2 x_2^u). \tag{9}$$

Since $X = P$, the quantities in equations (8) and (9) are equal. However, from Section 2.2, the $\mathbf{x}^{\mathrm{OPT}}$ which solves the LP in equation (8) is known and we can readily identify a u^{OPT} giving rise to this performance. When $c_1\mu_1 \geqslant c_2\mu_2$, $\mathbf{x}^{\mathrm{OPT}} = A$ and a control achieving this is $1 \to 2$. When $c_2\mu_2 \geqslant c_1\mu_1$, $\mathbf{x}^{\mathrm{OPT}} = B$ and this is achieved by $2 \to 1$. We thus conclude that the control solving equation (9) is the so-called *$c\mu$-rule* which gives priority to the customer class with the larger $c_k\mu_k$-value. Hence the optimal control favours options which drive down the holding cost rate most rapidly.

For certain classes of system, the above analysis can be generalized as follows: equation (4) is replaced by a *generalized work conservation law* for the entire set E of customer classes, given by

$$\sum_{j \in E} V_j^E x_j^u = b(E), \qquad u \in \mathcal{U}, \tag{10}$$

with $V_j^E, j \in E$, a set of positive constants. To generalize inequalities (5) and (6) suitably, we must consider an arbitrary subset S of customer classes. We then have

$$\sum_{j \in S} V_j^S x_j^u \geqslant b(S), \qquad u \in \mathcal{U}, \tag{11}$$

for positive constants $V_j^S, j \in S$, with the right-hand side of inequality (11) attained by any control which gives customers in S priority over those not in S. This latter requirement is expressed by

$$\sum_{j \in S} V_j^S x_j^u = b(S) \qquad \text{for } u: S \to S^c \tag{12}$$

Note that the approach requires that expressions (11) and (12) hold for all proper subsets of E. Bertsimas and Niño-Mora (1996) referred to expressions (10)–(12) as *generalized conservation laws* (GCLs). It follows trivially from these laws that when performances are positive valued the performance space X must be contained in the convex polyhedron

$$P = \left\{ \mathbf{x} \in (\Re^+)^N; \sum_{j \in S} V_j^S x_j \geqslant b(S),\ S \subset E, \text{ and } \sum_{j \in E} V_j^E x_j = b(E) \right\}. \tag{13}$$

Bertsimas and Niño-Mora (1996) further showed that when X is convex it must be true that $P \subseteq X$. They demonstrated that each extreme point of P must be the performance associated with a static priority policy, namely a control which chooses between customers according to a fixed ordering of their classes. Each non-extreme point can be expressed as a convex combination of extreme points and hence is easily seen to be the performance associated with a suitable randomization of a collection of such policies. It follows that $X = P$.

We suppose that the stochastic optimization problem of interest can be expressed as

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} \left(\sum_{j \in E} c_j x_j^u \right) = \min_{\mathbf{x} \in P} \left(\sum_{j \in E} c_j x_j \right). \tag{14}$$

Now, the LP on the right-hand side of equation (14) can be shown to be solved by the performance $\mathbf{x} = \mathbf{x}^{u_{\mathrm{G}}}$ of a *Gittins index policy* and hence, by an argument that is similar to that used in our simple example, such a control must solve the stochastic optimization problem. The Gittins index policy u_{G} operates by giving each customer class k an index G_k and then implementing priorities among the customer classes in E according to these indices, with the maximal index class being accorded highest priority. The indices are obtained from the so-called adaptive greedy algorithm $\mathrm{AG}(V, \mathbf{c})$ whose inputs are the matrix $V = \{V_j^S, j \in S, S \subseteq E\}$ and the cost vector $\mathbf{c}$. See Glazebrook and Niño-Mora (1997) for a description of the algorithm which computes the indices in order from smallest to largest. From this approach, Gittins index policies can be shown to be optimal for a reasonably general class of single-server models for the dynamic scheduling of stochastic projects and multiclass queues called branching bandits. See Weiss (1988) for details. These models include many classical ones, including the discounted multiarmed bandit of Gittins (1979, 1989) and the multiclass queue with Bernoulli feed-back of Klimov (1974). Systems whose optimal control can be characterized by a set of priority indices as described above are called *indexable* in Bertsimas and Niño-Mora (1996). In that sense, a linear objective together with GCLs (10)–(12) are sufficient for indexability. See Whittle (1988) for a different notion of indexability developed in the context of an analysis of restless bandits.

In the above examples of stochastic optimization problems based on GCL systems the performance space X can be identified (exactly). When this is not possible we may be able to proceed as follows. Firstly, identify constraints (assumed linear for our purposes) that are satisfied by the performance of choice $\mathbf{x}^u$ under all $u \in \mathcal{U}$. Such constraints will determine P, a subset of $(\Re^+)^N$ containing X. Assuming a linear objective, equation (14) is now replaced by

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} \left(\sum_{j \in E} c_j x_j^u \right) \geqslant \min_{\mathbf{x} \in P} \left(\sum_{j \in E} c_j x_j \right)$$

where the right-hand optimization problem is an LP with associated value Z^{LP}. Consider the dual of this LP and suppose that we can identify a feasible solution with value Z^D. Writing Z^u for the cost associated with control u, we invoke standard LP theory (weak duality) to infer that

$$Z^u \geqslant Z^{\mathrm{OPT}} \geqslant Z^{\mathrm{LP}} \geqslant Z^D. \tag{15}$$

Inequalities (15) can be used in (at least) two different ways to shed light on the performance of some heuristic control u of interest. Firstly, a computational approach could evaluate Z^{LP} as a lower bound on the minimized cost and use it as a yardstick to assess control u. Secondly, from inequalities (15), the quantity $Z^u - Z^{\mathrm{OPT}}$ which measures the degree of (cost) sub-optimality of control u is bounded above by $Z^u - Z^D$. It is sometimes possible to make choices of u and a feasible solution to the above dual such that this latter quantity (or upper bounds on it) may be identified as a helpful performance guarantee on u. Inequality (35) in Section 3 is an example of what can be achieved.

It is such ideas which lie behind the recent contributions of Glazebrook and Garbe (1999) and Glazebrook and Niño-Mora (1997) which were prompted by the discovery that many systems of interest come close to satisfying the GCL requirements (10)–(12) above but fail to do so exactly. In such cases, Gittins index policies may reasonably be expected to perform well for linear objectives, if not optimally. It is possible to use these ideas based on the primal–dual structure of an LP to obtain performance guarantees for index policies. In Section 3 this methodology is exploited to develop an analysis of a general multiclass queuing system serviced by M servers in parallel. In the single-server case $M = 1$, expressions (10)–(12) are satisfied exactly and Gittins index policies are optimal for a linear objective. When $M > 1$, we can develop measures of how close we come to achieving this (in theorem 1) which in turn leads (in corollary 1) to performance guarantees for such policies.

Another avenue of recent development has concerned work aimed at developing our understanding of how the optimal return Z^{OPT}depends on the mix of customer classes requiring service. Garbe and Glazebrook (1998a) investigated system properties which yield laws of diminishing returns (increasing marginal costs) as more demands are placed on the system. This work is exploited in Section 4 to develop an approach to distributing the load across a network of interconnected stations, when the work offered at each station is itself to be scheduled optimally. Note that the work in Sections 3 and 4 is in part motivated by a desire to consider systems of greater complexity than the single-server models of classical index theory. This seems an important research priority, particularly in the light of (for example) recent advances in parallel computing on multiprocessor systems.

We finally pause to note that the achievable region approach has recently found application outside the scope of the GCL framework. Bertsimas (1995) discusses polling systems, multiclass queuing networks and loss systems. Niño-Mora (1998) has begun a study of intensity control problems from this perspective. Some early conclusions are presented in Section 5.

3. A general multiclass queue on parallel servers

We consider here the optimal control of a general M-server queuing system. In the single-server case $M = 1$, we demonstrate that the system satisfies GCL (10)–(12) and in consequence Gittins index policies are optimal for a linear objective. The analysis of this case will show the reader how GCLs may be established in practice. Note that this index result is not new. See Bertsimas *et al.* (1995) for an account. What is new here is our development of that analysis to tackle the notoriously difficult parallel server problem with $M > 1$. Here we do not have exact GCLs but we come close. As a consequence, Gittins index policies come close to optimality. Following the work of Glazebrook and Garbe (1999), the achievable region approach furnishes us with performance guarantees for index policies, from which their asymptotic optimality in a heavy traffic limit may be inferred.

M servers are available to process the requirements of customers from classes in $E = \{1, 2, \ldots, N\}$. An assignment of available customers to servers is made at each integer time point. If a class i customer is assigned to server m at time t (which occurrence is registered by assigning the indicator function $I_i^m(t)$ the value 1; it is 0 otherwise) then at time $t + 1$ that class i customer disappears to be replaced by $\mathbf{n}_i^{tm} \equiv \{n_{i1}^{tm}, n_{12}^{tm}, \ldots, n_{iN}^{tm}\}$ customers of classes 1, 2, . . ., N respectively. For a given $i \in E$, the vectors $\mathbf{n}_i^{tm}$ are independent and identically distributed (IID) as (t, m) varies and for simplicity t (and sometimes also m) will be dropped from the notation when no confusion arises. As we shall see, this modelling approach enables us to incorporate state transitions for existing customers as well as new arrivals into the system. To complete the description of the system, note that an idle server is deemed to be serving a class 0 customer and we suppose that there are always M such customers in the system, one for each server. This additional class is needed to ensure that the model allows new arrivals to enter an empty system. We extend the notation $\mathbf{n}_i^{tm}$ to include the case $i = 0$, but note that $n_{00} = 1$ and $n_{i0} = 0$ for all $i \neq 0$.

If $N_i(t)$ denotes the number of class i customers in the system at decision epoch t, then the evolution of the system between t and $t+1$ is described by

$$N_i(t+1) = N_i(t) + \sum_{m=1}^{M} \sum_{j=0}^{N} I_j^m(t)(n_{ji}^m - \delta_{ij}), \qquad i \in E, \tag{16}$$

$$N_0(t+1) = N_0(t) = M.$$

In equation (16), δ_{ij} is the Kronecker delta. The set of admissible controls $\mathcal{U}$ available are

(a) non-anticipative,
(b) non-idling (which here means that E has priority over 0) and
(c) server symmetric (scheduling systems do not use server label information).

This third requirement is not strictly needed. It has been included for simplifying the discussion at certain key points. We can guarantee the *stability* of this system under all $u \in \mathcal{U}$ (the time-average number of customers in the system is finite) by requiring that the $N \times N$ matrix $\mathbf{I} - \mathbf{n}$ be positive definite. Here $\mathbf{I}$ is the identity and $\mathbf{n}$ has (i, j)th entry equal to $E(n_{ij})$. See Bertsimas and Niño-Mora (1996). We shall assume that admissible controls result in a discrete time stochastic process $\{N(t)\}_{t=-\infty}^{\infty}$ with unique stationary distribution, all of whose second moments are finite. Write

$$\rho_i^u = E_u\{I_i^m(t)\}, \qquad i \in E \cup \{0\}, \tag{17}$$

for the probability that control u assigns server m to a class i customer at decision epoch t, where the expectation in equation (17) is with respect to the stationary distribution. That this expectation does not depend on t (by stationarity) and m (by server symmetry) is clear. However, it turns out that it is also independent of the control u. To see this, apply E_u to both sides of equation (16) and use

$$E_u\{N_i(t+1)\} = E_u\{N_i(t)\}, \qquad i \in E,$$

to infer that $\boldsymbol{\rho}^u$ satisfies the system of equations

$$\sum_{i=0}^{N} \rho_i^u \, E(n_{ij}) = \rho_j^u, \qquad j \in E,$$

$$\sum_{i=0}^{N} \rho_i^u = 1. \tag{18}$$

This has a unique solution when $\mathbf{I} - \mathbf{n}$ is non-singular. We shall write ρ without the superscript in what follows. One particular focus of the analysis will concern the quality of the control policies in *heavy traffic*. In discussing a sequence of systems, we are said to approach the heavy traffic limit if the value ρ_0 (the steady state probability that a server is idle) approaches 0.

3.1. Example

Consider a discrete time version of a multiclass $M/G/parallel$ queuing system with M servers and customer feed-back as follows: customers belonging to one of L classes arrive for service according to independent Poisson streams with λ_l the rate for class l, $1 \leqslant l \leqslant L$. Service times T_l for class l customers are IID discrete random variables whose distribution has finite support $\{1, 2, \ldots, R_l\}$. F_l denotes the distribution function of T_l. Once a class l customer has completed service, she or he is fed back to the system as a class k customer with probability p_{lk} or leaves the system with probability $p_{l0} = 1 - \Sigma_{k=1}^{L} p_{lk}$. The scheduling regime gives to each customer chosen for service a single unit of processing before the position is reviewed again.

It is straightforward to cast this example into the general framework above. We require classes labelled $\{(l, r), 0 \leqslant r \leqslant R_l - 1, 1 \leqslant l \leqslant L\}$ with (l, r) representing those class l customers in the system who have already received r units of processing. A newly arrived class l customer (either from outside or via feed-back) belongs to $(l, 0)$. When a unit of processing is allocated to a class (l, r) customer, there are two possibilities: either there is a failure to complete service and the customer is now in class $(l, r+1)$ or service is completed and the customer leaves the system or feeds back as a $(k, 0)$ customer for some k. This, together with consideration of external arrivals, yields the following choices of the components of the matrix $\mathbf{n}$:

$$
\begin{aligned}
E\{n_{(l,r),\,(l,r+1)}\} &= \frac{1 - F_l(r+1)}{1 - F_l(r)}, && 0 \leqslant r \leqslant R_l - 1, \quad 1 \leqslant l \leqslant L, \\
E\{n_{(l,r),\,(k,0)}\} &= \frac{\lambda_k}{M} + \frac{\{F_l(r+1) - F_l(r)\} p_{lk}}{1 - F_l(r)}, && 0 \leqslant r \leqslant R_l - 1, \quad 1 \leqslant l \leqslant L, \quad 1 \leqslant k \leqslant L, \\
E\{n_{0,\,(k,0)}\} &= \frac{\lambda_k}{M}, && 1 \leqslant k \leqslant L.
\end{aligned} \tag{19}
$$

Now introduce the system parameters α_l, $1 \leqslant l \leqslant L$, obtained as the solution to the linear system

$$
\alpha_l = \lambda_l + \sum_{k=1}^{L} \alpha_k p_{kl}, \qquad 1 \leqslant l \leqslant L. \tag{20}
$$

The quantity α_l is easily seen to be the *total arrival rate* for class l customers, aggregating the external arrival rate (λ_l) with an internal rate (the second term on the right-hand side of equation (20)) obtained via feed-back. Substituting from equations (19) and (20) into the equations (18) we obtain solutions in the form

$$
\rho_{(l,r)} = \alpha_l \frac{1 - F_l(r)}{M}, \qquad 0 \leqslant r \leqslant R_l - 1, \quad 1 \leqslant l \leqslant L.
$$

Hence we deduce that

$$
\rho_0 = 1 - \sum_{l=1}^{L} \sum_{r=0}^{R_l - 1} \rho_{(l,\,r)} = 1 - \sum_{l=1}^{L} \alpha_l \frac{E(T_l)}{M} \rightarrow 0 \tag{21}
$$

in the heavy traffic limit. Note that $\Sigma_l \, \alpha_l \, E(T_l)$ measures the rate at which work is created. Hence equation (21) asserts that the heavy traffic limit is attained as this approaches M, the total service rate available.

Returning to our general system, our stochastic optimization problem is expressed as

$$Z^{\mathrm{OPT}} = \inf_{u \in \mathcal{U}} \left(\sum_{i \in E} c_i x_i^u \right) \tag{22}$$

where $c_i > 0$, $i \in E$, and

$$x_i^u = E_u\{N_i(t)\}, \qquad i \in E, \tag{23}$$

the expectation being taken under the stationary distribution. We would like to generate linear (in)equalities of the forms (10) and (11). For this, we deploy the potential function approach of Bertsimas *et al.* (1995) who considered the quantity $R^S(t)^2$, where

$$R^S(t) = \sum_{i \in S} V_i^S N_i(t), \tag{24}$$

for suitably chosen V_i^S, $i \in S$. Note from equation (23) that $E_u\{R^S(t)\}$ is precisely the quantity on the left-hand side of inequality (11). We choose the positive constants V_i^S, $i \in E$, as solutions of the linear system

$$V_i^S = 1 + \sum_{j \in S} E(n_{ij}) V_j^S, \qquad i \in E. \tag{25}$$

Observe that V_i^S, $i \in S$, may be thought of as the mean amount of S-work required by a class i customer — i.e., beginning from a situation in which only a single class i customer is present and under a control which gives priority to S, this is the mean amount of processing required until no S-customers are present. Hence $R^S(t)$ is the total amount of S-work in the system at t. Using equations (16) and (24) we infer that

$$R^S(t+1) = R^S(t) + \sum_{i \in S} \sum_{m=1}^{M} \sum_{j=0}^{N} V_i^S I_j^m(t)(n_{ji}^m - \delta_{ij}). \tag{26}$$

If we square both sides of equation (26), take E_u and enforce the stationarity condition

$$E_u\{R^S(t+1)^2\} = E_u\{R^S(t)^2\}$$

we infer the condition

$$2\,E_u\left\{ R^S(t) \sum_{m=1}^{M} \sum_{j=0}^{N} I_j^m(t) \sum_{i \in S} V_i^S (n_{ji}^m - \delta_{ij}) \right\} + E_u\left[\left\{ \sum_{m=1}^{M} \sum_{j=0}^{N} I_j^m(t) \sum_{i \in S} V_i^S (n_{ji}^m - \delta_{ij}) \right\}^2 \right] = 0. \tag{27}$$

Considerable simplification of equation (27) is possible which exploits the server symmetry of u and the fact that $I_j^m(t)\, I_k^m(t) = 0$ whenever $j \neq k$. Straightforward algebra yields

$$\begin{aligned}\sum_{i \in S} V_i^S x_i^u &= E_u\{R^S(t)\} \\ &= E_u\left\{ R^S(t) \sum_{i \notin S} V_i^S I_i^1(t) \right\} + b(S) + \frac{1}{2}(M-1)\,E_u\left[\left\{ \sum_{i \notin S} V_i^S I_i^1(t) \right\} \left\{ \sum_{j \notin S} V_j^S I_j^2(t) \right\} \right],\end{aligned} \tag{28}$$

where $b(S)$ is a control invariant constant given by

$$b(S) = \frac{1}{2} \sum_{j=0}^{N} \rho_j\, E\left[\left\{ \sum_{i \in S} V_i^S (n_{ji} - \delta_{ij}) \right\}^2 \right] + \frac{1}{2}(M-1)\left(1 - 2 \sum_{i \notin S} \rho_i V_i^S \right). \tag{29}$$

It is from equation (28) that we can develop suitable forms of the (in)equalities (10) and (11) for this system. The requirement in equation (12) which we need for a full GCL–Gittins index analysis as described in Section 2 is satisfied in the single-server case $M = 1$. When $M > 1$ we come close to having requirement (12) in a sense which is made precise in the following result. Before stating it, note that equation (10) may be regarded as a particular case of equation (12), namely for $S = E$. We shall also require the notation $b^+ = \max(b, 0)$ for set functions b.

Theorem 1 (exact and approximate GCL for the system).

(a) For all values of M and all controls $u \in \mathcal{U}$

$$\sum_{i \in S} V_i^S x_i^u \geqslant b^+(S), \qquad S \subseteq E. \tag{30}$$

(b) In the single-server case $M = 1$ the system satisfies GCLs, i.e. in addition to (a) we have

$$\sum_{i \in S} V_i^S x_i^u = b^+(S) = b(S) \qquad \text{for } u\text{: } S \rightarrow S^c, \quad S \subseteq E.$$

(c) When there is more than one server, $M \geqslant 2$, controls u which give S priority over S^c come within a finite constant of achieving the bound on the right-hand side of inequality (30). In particular

$$\sum_{i \in S} V_i^S x_i^u \leqslant b^+(S) + \frac{1}{2}(M-1)(3+\hat{n})V^S, \qquad S \subseteq E, \tag{31}$$

where

$$\hat{n} = \max_{i \in E} \left\{ \sum_{j \in E} E(n_{ij}) \right\} \quad \text{and} \quad V^S = \max_{i \in S}(V_i^S).$$

Outline proof.

(a) The left-hand side of inequality (30) must be non-negative as must the first and third terms on the right-hand side of equation (28). Part (a) is then an immediate consequence of equation (28).
(b) If $M = 1$ and u gives S priority over S^c then, under u,

$$R^S(t) > 0 \Longrightarrow N_i(t) > 0 \quad \text{for some } i \in S \Longrightarrow I_i^1(t) = 0, \quad i \notin S.$$

Hence the first term on the right-hand side of equation (28) is 0. Since the third term is 0 trivially, the result follows.
(c) In the case $M \geqslant 2$ we are required to bound the first and third terms in equation (28) from above for policies u which give S priority over S^c. Taking the first term, we can assert that, since under u

$$\sum_{i \in S} N_i(t) \geqslant M \Longrightarrow I_i^1(t) = 0, \qquad i \notin S,$$

it follows that

$$E_u\left\{R^S(t)\sum_{j\notin S} V_j^S I_j^1(t)\right\} \leqslant (M-1)V^S \sum_{j\notin S} V_j^S E_u\{I_j^1(t)\}$$
$$= (M-1)V^S \sum_{j\notin S} \rho_j V_j^S = (M-1)V^S. \quad (32)$$

In expression (32), note that $\Sigma_{j\notin S}\,\rho_j V_j^S = 1$ may be established *either* algebraically (from equations (18) and (25)) *or* by the use of probabilistic arguments. We now consider the third term in equation (28). A Cauchy–Schwarz inequality yields

$$E_u\left[\left\{\sum_{i\in S} V_i^S I_i^1(t)\right\}\left\{\sum_{i\notin S} V_j^S I_j^2(t)\right\}\right] \leqslant E_u\left[\left\{\sum_{i\notin S} V_i^S I_i^1(t)\right\}^2\right]$$
$$= \sum_{j\notin S} \rho_j (V_j^S)^2 \leqslant (1+\hat{n})V^S. \quad (33)$$

The last inequality in expression (33) follows simply from equation (25). The result is now a straightforward consequence of expressions (28), (32) and (33). □

We see from theorem 1, parts (a) and (b), and the material in Section 2 that in the single-server case $M = 1$ the requirements described in expressions (10)–(12) are met (i.e. GCLs are satisfied) and the stochastic optimization problem (22) is solved by a Gittins index policy. The indices concerned are derived from the adaptive greedy algorithm AG(V, **c**).

In the parallel server case with $M \geqslant 2$ we proceed as follows: from theorem 1, part (c), the set function Φ given by

$$\Phi(S) = \tfrac{1}{2}(M-1)(3+\hat{n})V^S, \qquad S \subseteq E, \quad (34)$$

is a natural measure of how close we come to satisfying the GCL requirement in equation (12). As sketched briefly in Section 2, Glazebrook and Garbe (1999) utilized the primal–dual structure of an LP to develop a performance guarantee for the Gittins index policy derived from AG(V, **c**) in terms of the measure Φ. Numerical and analytical evidence to date suggests that the tightest such guarantees available perform very well in bounding the level of sub-optimality of the index policy u_G. We shall give a somewhat simplified account here which will be sufficient for our purposes. Note that the bounds that we shall describe are by no means the tightest that are available from the methodology.

Application of AG(V, **c**) yields the indices G_i, $i \in E$. The customer classes are then renumbered such that $G_N \geqslant G_{N-1} \geqslant \ldots \geqslant G_1$. Hence, the index policy u_G implements priorities among the customer classes in decreasing numerical order. We identify $S(j) = \{j, j+1, \ldots, N\}$ as the set of cardinality $N-j+1$ of classes with highest index. Note that u_G prefers $S(j)$ to $S(j)^c$ for all j. Our goal here is to develop a bound for $Z^{u_G} - Z^{\mathrm{OPT}}$ where Z^{u_G} is the expected cost associated with the Gittins index policy. From the work of Glazebrook and Garbe (1999) we have that

$$Z^{u_G} - Z^{\mathrm{OPT}} \leqslant \sum_{j=1}^{N} \Phi\{S(j)\}(G_j - G_{j-1}) \quad (35)$$

where Φ is the above measure and G_0 is taken to be 0.

It is not difficult now to establish corollary 1 by substituting from equation (34) into inequality (35) and utilizing the form of the algorithm AG(V, **c**) which produces the indices.

Corollary 1 (performance guarantee for Gittins index policy when $M \geqslant 2$).

$$Z^{u_G} - Z^{OPT} \leqslant \tfrac{1}{2} N(M-1)(3+\hat{n}) \max_{i \in E}(c_i).$$

One remarkable thing about the claim in corollary 1 that the index policy comes within a constant of optimality is that the optimum cost Z^{OPT} becomes infinite (under reasonable conditions) as the heavy traffic limit is approached. Hence u_G is asymptotically optimal in a sense made precise below. Such a result is not unexpected. Index policies are optimal in the single-server case since they always make choices which drive down the rate at which costs are incurred as rapidly as possible. The parallel server case is complicated by the issue of how effectively controls utilize the full service capacity. (Attempts to tackle these issues directly have met little success. See Weiss (1992, 1995) for an authoritative discussion in the context of much simpler models than those cited here.) However, and to oversimplify the issues concerned hugely, in the heavy traffic limit server utilization disappears as an issue and the system looks increasingly like one serviced by a single server working at M times the speed.

To establish the asymptotic optimality of u_G we can infer from inequality (30) with $S = E$ that

$$Z^{OPT} = \sum_{i \in E} c_i x_i^{OPT} \geqslant \min_{j \in E}(c_j/V_j^E) \sum_{i \in E} V_i^E x_i^{OPT} \geqslant b(E) \min_{j \in E}(c_j/V_j^E) \tag{36}$$

with the set function b given by equation (29). It will be enough to investigate conditions which guarantee that the right-hand side of expression (36) diverges to ∞ in the heavy traffic limit $\rho_0 \to 0$. One way of achieving this is as follows: suppose that the vectors $\mathbf{n}_i$ record two types of change to the composition of the system, namely

(a) external arrivals into customer classes within some designated subset $A \subseteq E$ and
(b) internal transfers via feed-back or some other transition mechanism.

Plainly, our example above of an $M/G/$parallel system with feed-back may be thought of in these terms. Hence when $i \in E \cup \{0\}$ we write

$$n_{ij} = \begin{cases} A_j + \tilde{n}_{ij}, & j \in A, \\ \tilde{n}_{ij}, & \text{otherwise.} \end{cases} \tag{37}$$

In equations (37), A_j denotes external arrivals to j (assumed independent of all other A_i and all the $\tilde{n}_{kl}$) and $\tilde{n}_{ij}$ internal transfers from i to j. We assume that $E(A_j) = \lambda_j/M$, where λ_j is an overall class j arrival rate for the system. We shall approach the heavy traffic limit by increasing the λ_j appropriately while

(a) keeping the $E(\tilde{n}_{ij})$ fixed and
(b) keeping $\text{var}(A_j)$ bounded away from 0.

Note that (b) is required to avoid certain pathologies which occur in deterministic cases. Note also that all this is quite natural in the $M/G/$parallel case.

Utilizing equations (37) within an expanded version of equation (25) which includes the 'idleness' class 0, we can solve for $\mathbf{V}^E = [V_j^E, j \in E \cup \{0\}]$, obtaining

$$\mathbf{V}^E = \hat{V}(\mathbf{I} - \tilde{\mathbf{n}})^{-1}\mathbf{e}, \tag{38}$$

where in equation (38)

$$\hat{V} = 1 + \sum_{j \in A} \lambda_j V_j^E / M,$$

e is a vector with all entries equal to 1 and $\tilde{\mathbf{n}}$ is a matrix whose (i, j)th entry is $E(\tilde{n}_{ij})$. Note that $\mathbf{I} - \tilde{\mathbf{n}}$ is guaranteed non-singular by earlier assumptions.

Recall from the proof of theorem 1 the identity $\Sigma_{j \notin S} \rho_j V_j^S = 1$. In the case $S = E$ this yields $\rho_0 V_0^E = 1$. Hence in the heavy traffic limit $\rho_0 \to 0$ and $V_0^E \to \infty$. However, in equation (38), since we have assumed that $\mathbf{I} - \tilde{\mathbf{n}}$ remains fixed as we take the limit, it must follow that $\hat{V} \to \infty$ and hence that $V_j^E \to \infty, j \in E$. We can now assert the asymptotic optimality of the Gittins index policy u_{G}.

Theorem 2 (heavy traffic optimality of Gittins index policy when $M \geqslant 2$). In the above heavy traffic limit

$$\frac{Z^{u_{\mathrm{G}}} - Z^{\mathrm{OPT}}}{Z^{\mathrm{OPT}}} \to 0.$$

Proof. We utilize equation (29) to obtain $b(E)$. By standard results and the fact that $\rho_0 V_0^E = 1$, we deduce that

$$2\,b(E) \geqslant \sum_{j=0}^{N} \rho_j \operatorname{var}\left(\sum_{i \in E} V_i^E n_{ji}\right) - (M-1)$$

$$\geqslant \sum_{j=0}^{N} \rho_j \operatorname{var}\left(\sum_{i \in A} V_i^E A_i\right) - (M-1) \tag{39}$$

$$= \sum_{i \in A} (V_i^E)^2 \operatorname{var}(A_i) - (M-1). \tag{40}$$

To obtain inequality (39), we use equations (37) and the independence assumptions following. From expressions (36), (38) and (40) we conclude that

$$Z^{\mathrm{OPT}} \geqslant O(\hat{V}) \to \infty$$

in the heavy traffic limit. The result now follows from corollary 1. □

4. Load balancing in distributed systems

A common architecture for multiprocessor systems is a distributed one consisting of a network of (relatively) autonomous servers. The issue of the efficient allocation of resources in such contexts is both important and complex. See Gelenbe and Pekergin (1993). One fundamental question concerns the distribution of work across the network or, as we shall call it, load balancing. The theoretical literature has, in the main, concentrated on very simple models. For these, simple round robin policies and Bernoulli routing with equal probabilities have frequently been proposed as optimal load balancing regimes when little information is available to the controller. See, for example, Liu and Towsley (1994). When full information on queue lengths is available, the join the shortest queue strategy has been shown to be optimal for a variety of models. See Weber (1978).

In a contribution which represented a significant advance, Ross and Yao (1991) showed that considerable savings could be made if optimal scheduling of the work offered at each station of the network could be incorporated into the load balancing problem. Their work made use of the achievable region approach, but predated many of the most significant advances outlined in Section 2. The authors of the current paper and co-workers plan a much more extensive study and we report here some of the early findings.

We shall consider a communication network interconnecting multiple stations, with two types of jobs generated at each station: those which are dedicated (D) to that station and must be processed there and those which are generic (G) and could be processed anywhere in the network. There may be several classes of D- and G-jobs, arriving in independent Poisson streams. We seek to split the G-traffic between the individual stations in an optimal fashion given that each station schedules its offered work (both D and G) optimally. On the basis of a realistic appraisal of the communication and processing overhead generated thereby our policies for scheduling at each station will be dynamic (i.e. decisions will be made on the basis of the evolving state of each station) whereas the load balancing component of the problem will be static (i.e. the vector of generic arrival rates will be split once for all between the stations). At this point we introduce two simple examples to assist the reader.

4.1. Examples

It may seem plausible to conjecture that when the stations in the network are identical (in all relevant respects) then an optimal load balancing regime will split the G-jobs equally between them. The following simple examples will caution the reader against drawing such conclusions too easily. In both examples the network comprises two identical single-server stations. In each case there are two G-job classes and no D-jobs. The objective in both examples is the minimization of $c_1 E(N_1) + c_2 E(N_2)$ where $E(N_i)$ is the expected number of class i jobs in the system and the expectation is taken under the steady state distribution of the corresponding stochastic process with c_i a holding cost rate, $i = 1, 2$.

4.1.1. *Example 1*

Here we shall suppose that generic job class 1 has zero holding costs ($c_1 = 0$) but a high arrival rate to the network ($\lambda_1 = 0.9$), whereas for job class 2 we have positive holding costs ($c_2 = 1$) and a low arrival rate ($\lambda_2 = 0.1$). The processing time of all jobs is exactly 1 and at each station scheduling is non-pre-emptive. Plainly at both stations the optimal scheduling regime prefers class 2 to class 1 and must impose that priority in a fashion that is non-pre-emptive.

Since $c_1 = 0$, our objective is to split the load in order to minimize $E(N_2)$. An even split of arriving jobs between the stations will result (frequently) in situations where an arriving class 2 job finds the machine busy with a class 1 job and is thus delayed while its processing is completed. An alternative regime in which all class 1 jobs go to one machine and all class 2 jobs go to the other will result in less frequent delays to the latter because λ_2 is small. Simple calculations show that the 'one job class per machine' regime yields a 16.43% saving in expected cost over an even distribution of work.

4.1.2. *Example 2*

Plainly the non-pre-emptive nature of the scheduling regime plays a significant role in example 1. Consider now a situation in which scheduling priorities are imposed pre-emptively. We shall suppose that all processing times are exponentially distributed with mean 1 for class 1 and mean 0.1 for class 2. The usual full range of independence assumptions is made. We also take $c_1 = 1$, $\lambda_1 = 0.5$, and $c_2 = 0.1$, $\lambda_2 = 10$. Direct calculations show that a (nearly optimal) splitting of the load in which all class 1 traffic is directed to station 1, while 85% of class 2 traffic goes to station 2, offers a 17.25% saving in expected cost over an even distribution of work. Note that example 2 elaborates example 1 in that processing times for generic jobs are not identically distributed. Subsequent theory serves to show that this is a required feature for an even split solution to be suboptimal with exponential processing times and priorities imposed pre-emptively.

We shall suppose that our load balancing problem may be expressed as

$$\min_{\Lambda}\left\{\sum_m Z_m^{\mathrm{OPT}}(\boldsymbol{\lambda}_m)\right\} \qquad \text{subject to } \boldsymbol{\Lambda}\mathbf{e}=\boldsymbol{\lambda}. \tag{41}$$

In problem (41), λ_{gm} is the offered load (i.e. arrival rate) of class g jobs at station m, where $g \in G$, $\boldsymbol{\lambda}_m$ is the vector of generic loads at station m and $\boldsymbol{\Lambda} = \{\lambda_{gm}\}$ is the generic load matrix. Vector $\boldsymbol{\lambda}$ summarizes the total generic load for the network and $\mathbf{e}$ is an M-vector of 1s, where M is the number of stations. $Z_m^{\mathrm{OPT}}(\boldsymbol{\lambda}_m)$ is the minimized cost at station m when $\boldsymbol{\lambda}_m$ is the generic load offered there. This minimized cost is achieved when the offered work is scheduled optimally.

Plainly, an ability to compute and/or characterize the returns Z_m^{OPT} as functions of the generic load vectors $\boldsymbol{\lambda}_m$ will contribute to achieving optimal or nearly optimal solutions to problem (41). As we shall now see, we can make considerable progress when each station satisfies GCLs. We drop the station suffix m as we carry the discussion forward regarding the individual stations in the network.

Happily it is one of the features of the GCL and indexable systems described in Section 2 that computations of expected cost for a given (priority) policy can be performed with ease, as can the computation of minimized cost Z^{OPT}. See Bertsimas and Niño-Mora (1996). In addition, Z^{OPT} can often be characterized in a way which will ultimately assist with problem (41) as follows. Consider a GCL system with linear costs and an associated universal set E of *potential* customer classes. For specified $S \subseteq E$, let $Z^{\mathrm{OPT}}(S)$ be the minimized cost for the reduced system in which only customer classes within S are allowed access to service. Garbe and Glazebrook (1998a) showed that the achievable region approach yields the conclusion that, subject to some additional structural requirements, Z^{OPT} is an *increasing and supermodular* function, namely

$$\begin{gathered} Z^{\mathrm{OPT}}(S) \leqslant Z^{\mathrm{OPT}}(T), \qquad S \subseteq T \text{ (increasing)}, \\ Z^{\mathrm{OPT}}(S \cup \{j\}) - Z^{\mathrm{OPT}}(S) \leqslant Z^{\mathrm{OPT}}(T \cup \{j\}) - Z^{\mathrm{OPT}}(T), \\ S \subseteq T \text{ and } j \notin T \text{ (supermodular)}. \end{gathered} \tag{42}$$

Supermodularity states, in this context, that allowing an additional class of customers access to a more congested system increases the optimum cost by more than allowing the same additional class access to a less congested system. This seems a natural property for Z^{OPT}.

We shall want to draw on this result in our discussion of load balancing. However, rather than to develop the theory through general model structures which have the properties required to establish expressions (42), for clarity we shall conduct the discussion in terms of a specific GCL model for each station which meets the requirements. Directions in which the material can be generalized are sketched at the end of the section.

4.2. Model for local scheduling at each station

We shall suppose that each station is a *Klimov network* (see Klimov (1974)) as follows (note that we continue to drop the station suffix m): customers who are members of classes within $D \cup G$ are assumed to arrive at the station in a set of independent Poisson streams. Use λ_g to denote an arrival rate for generic class $g \in G$ and $\boldsymbol{\lambda}$ the corresponding vector of generic arrival rates. All customers have exponential service times with mean denoted μ_g^{-1} for class $g \in G$. On completion of service, a class i customer may be routed to receive further service as a class j customer with probability p_{ij}, or it may leave the station with probability $p_{i0} = 1 - \Sigma_{j \in D \cup G} \, p_{ij}$. All the customer arrival processes, service times and routing events are mutually independent. The routing probability matrix $\mathbf{P} = (p_{ij}, \, i, j \in D \cup G)$ is such that $\mathbf{I} - \mathbf{P}$ is invertible, thus guaranteeing that a customer entering the system will leave it with probability 1. Under this model, the traffic at a station can be quite general in its structure. For example, the framework proposed allows customers to have a state which evolves in continuous time as a (finite state) Markov process through to completion. Admissible scheduling controls at the station are non-anticipative, non-idling and pre-emptive.

Each customer class $i \in D \cup G$ has an associated holding cost rate c_i and so the minimized cost Z^{OPT} for the station is given by

$$Z^{\text{OPT}}(\boldsymbol{\lambda}) = \inf_{u \in \mathcal{U}} \left(\sum_{i \in D \cup G} c_i x_i^u \right)$$

with $x_i^u = E_u(N_i)$, the long-term average number of class i customers at the station. From the theory described in Section 2, the control which achieves Z^{OPT} will be an index policy.

We now seek to characterize $Z^{\text{OPT}}(\boldsymbol{\lambda})$ as a function of the offered generic load $\boldsymbol{\lambda}$ (holding other model parameters fixed) over those $\boldsymbol{\lambda}$ which yield a stable system. For the efficient solution of problem (41), we would ideally like each $Z^{\text{OPT}}(\boldsymbol{\lambda})$ to be increasing and convex. However, in higher dimensions, full convexity is a very strong property and in general we must settle for the weaker form described in definition 1 in which convexity is available in certain directions only in load space.

Definition 1. A real-valued function f on generic load space is *north-east (NE) convex* if, for all $\alpha \in [0, 1]$ and all $\boldsymbol{\lambda}'$ and $\boldsymbol{\lambda}''$ such that $\boldsymbol{\lambda}' - \boldsymbol{\lambda}'' \geqslant 0$,

$$\alpha f(\boldsymbol{\lambda}') + (1 - \alpha) f(\boldsymbol{\lambda}'') \geqslant f\{\alpha \boldsymbol{\lambda}' + (1 - \alpha) \boldsymbol{\lambda}''\}.$$

Such directional convexity plays an important role in other areas of stochastic scheduling. See Chang and Yao (1993). There is a simple proof of theorem 3 which begins with the supermodularity property in expressions (42) and infers from that properties of $Z^{\text{OPT}}(\boldsymbol{\lambda})$. The argument essentially secures increased generic arrival rates through the introduction of new generic job classes with appropriate service characteristics. We omit the details.

Theorem 3. For our Klimov network model, the minimized cost Z^{OPT} is increasing and NE convex.

Note that NE convexity certainly includes convexity in each co-ordinate direction (for fixed values elsewhere). Further, in the one-dimensional case it coincides with full convexity. Corollary 2 follows.

Corollary 2. If $|G| = 1$, the minimized cost Z^{OPT} is increasing and convex.

Hence we have full convexity for the case of a single generic class. This can be readily extended in two directions. The first concerns situations in which the controller of the distributed system cannot distinguish between generic jobs. In this case, the solution to problem (41) will be of the form $\{\alpha_m, \, 1 \leqslant m \leqslant M\}$ where α_m is the proportion of all generic traffic passed to station m. Let $\boldsymbol{\lambda}$ now stand, as in expression (41), for the generic load for the network. The optimization goal becomes the minimization of

$$\sum_m Z_m^{\text{OPT}}(\alpha_m \boldsymbol{\lambda}) \qquad \text{subject to } \alpha_m \geqslant 0, \quad 1 \leqslant m \leqslant M, \quad \sum_m \alpha_m = 1. \tag{43}$$

To solve problem (43), our interest is in $Z^{\text{OPT}}(\alpha\boldsymbol{\lambda})$ as a function of α for fixed $\boldsymbol{\lambda}$, where $\alpha \in [0, 1]$. The following is an immediate consequence of theorem 3.

Corollary 3. For fixed $\boldsymbol{\lambda}$, the minimized cost $Z^{\text{OPT}}(\cdot\boldsymbol{\lambda})$: $\alpha \to Z^{\text{OPT}}(\alpha\boldsymbol{\lambda})$ is increasing and convex.

Another direction in which theorem 3 can be extended is to cover situations in which $|G| > 1$, but where the generic traffic is particularly simple in structure. We shall require that there be no generic feed-back, i.e.

$$p_{ij} = 0 \text{ for } i \in D, j \in G \text{ and } i \in G, j \in D \cup G,$$

and that generic processing requirements are IID.

Theorem 4. If $\mu_g = \mu$, $g \in G$, and there is no generic feed-back then the minimized cost Z^{OPT} is increasing and convex.

Outline proof. Write $N = |D \cup G|$. As in Section 3, the customer classes at our single station are renumbered such that $G_N \geqslant G_{N-1} \geqslant \ldots \geqslant G_1$ and again we write $S(j) = \{j, j+1, \ldots, N\}$. In the notation established in expressions (10)–(14) in Section 2, it will assist to express the dependence of the base function b on the generic arrival rate $\boldsymbol{\lambda}$. Hence we write $b(S, \boldsymbol{\lambda})$, $S \subseteq E$. Recall that Z^{OPT} is the value of the LP (14) and hence also of its dual. The latter was shown by Bertsimas and Niño-Mora (1996) to be expressible as

$$Z^{\text{OPT}}(\boldsymbol{\lambda}) = \sum_{j=1}^{N} b\{S(j), \boldsymbol{\lambda}\}(G_j - G_{j-1}) \tag{44}$$

where, in equation (44), $G_0 = 0$. Note also that it is straightforward to show for our Klimov network model that the indices G_j do not depend on $\boldsymbol{\lambda}$.

From the GCLs (11) and (12) we may write

$$\inf_{u \in \mathcal{U}} \left(\sum_{i \in S(j)} V_i^{S(j)} x_i^u \right) = b\{S(j), \boldsymbol{\lambda}\}. \tag{45}$$

It can be shown that, since the generic classes in $G \cap S(j)$ have IID processing requirements and there is no generic feed-back, then they must all have the same associated value of $V_g^{S(j)}$. Hence, they may be regarded corporately as a single customer class with arrival rate $\Sigma_{g \in G \cap S(j)} \lambda_g$ so far as the stochastic optimization problem on the left-hand side of equation (45) is concerned. It then follows from corollary 2 that we may write

$$b\{S(j), \boldsymbol{\lambda}\} \equiv b_j \left(\sum_{g \in G \cap S(j)} \lambda_g \right)$$

where b_j is increasing and convex. Theorem 4 now follows from equation (44), the $\boldsymbol{\lambda}$-independence of the indices and from basic properties of convex functions. □

We have established a range of scenarios in which the minimized cost at each station is increasing and convex in the generic load (corollaries 2 and 3 and theorem 4, with more to come) and a greater range for which convexity is available for certain directions in load space (including co-ordinatewise). We now consider briefly the implications for the load balancing problem (41). We begin by a consideration of the special case in which all stations are identical, i.e. the minimized cost for station m, $Z_m^{\mathrm{OPT}}(\cdot) \equiv Z^{\mathrm{OPT}}(\cdot)$, $1 \leqslant m \leqslant M$.

Theorem 5. When stations are identical, it is optimal to split the generic load evenly between stations for all loads $\boldsymbol{\lambda}$ if and only if $Z^{\mathrm{OPT}}(\cdot)$ is convex.

Proof. If $Z^{\mathrm{OPT}}(\cdot)$ is convex and $\boldsymbol{\lambda}_m$ is the generic load for station m as in problem (41), then

$$\sum_{m=1}^{M} Z^{\mathrm{OPT}}(\boldsymbol{\lambda}_m) \geqslant M\, Z^{\mathrm{OPT}}\left(\sum_{m=1}^{M} \frac{1}{M}\boldsymbol{\lambda}_m\right) = M\, Z^{\mathrm{OPT}}\left(\frac{1}{M}\boldsymbol{\lambda}\right), \tag{46}$$

by convexity. However, the final term in expression (46) is plainly the cost corresponding to an even load distribution. For the converse, see Dacre and Glazebrook (1999). □

It is possible to supplement theorem 5 via the development of performance guarantees for an even split of the generic load when full convexity for $Z^{\mathrm{OPT}}(\cdot)$ is not available. For example, if we take one of the simplest cases of interest, namely of two identical stations each having $|D| = 0$ and $|G| = 2$ and with no feed-back, then it can be shown that an even load distribution yields a cost which is within a fraction

$$\frac{1}{2}\,\frac{|\mu_1 - \mu_2|}{\mu_1 + \mu_2} \tag{47}$$

of the optimal cost for the network. See Dacre and Glazebrook (1999). Expression (47) is 0 when $\mu_1 = \mu_2$, indicating that an even distribution is optimal in the IID case with no feed-back. This is in agreement with theorems 4 and 5.

Note that, following theorem 3 and extensive numerical investigation, there are many systems for which the optimum cost, although not fully convex, comes close to being so. When the $Z_m^{\mathrm{OPT}}(\cdot)$ are indeed all convex, an efficient iterative procedure is available for the load balancing problem (41) which solves a sequence of LPs determined via subgradients of the objective. Our numerical study has shown that, in practice, this approach yields acceptable solutions even in the absence of full convexity. In Figs 2 and 3 we illustrate the performances of

(a) an even load distribution and
(b) this LP-based heuristic

for problem (41) for the simple case above, namely of two identical stations with $|D| = 0$ and $|G| = 2$ and no feed-back. Figs 2 and 3 are based on a grid of $(60)^2$ points with both $\log_e(c_2/c_1)$ and $\log_e(\mu_2/\mu_1)$ taken to be in the range from -3 to 3 in steps of 0.1. The values of μ_1 and c_1 are both set equal to 1, although the results presented are the same for any assigned values of these constants. At each grid point is presented a summary of the performance for the chosen load balancing regimes over 120 problems—each corresponding to a choice of generic arrival rate $\boldsymbol{\lambda}$. In Fig. 2, the chosen performance measure is the maximum percentage level of suboptimality of the load balancing regime over the 120 problems whereas in Fig. 3 we report the percentage of solutions which were within 0.01% of optimality. In interpreting the results, note the following.

(a) By theorem 5, we should expect the performance of the even load distribution to give an indication of the extent of non-convexity of Z^{OPT}.
(b) The formula for Z^{OPT} in equation (44) applied to the present case may be written

$$Z^{\mathrm{OPT}}(\boldsymbol{\lambda}) = |c_1\mu_1 - c_2\mu_2|\phi(\boldsymbol{\lambda}) + \min(c_i\mu_i)\psi(\boldsymbol{\lambda})$$

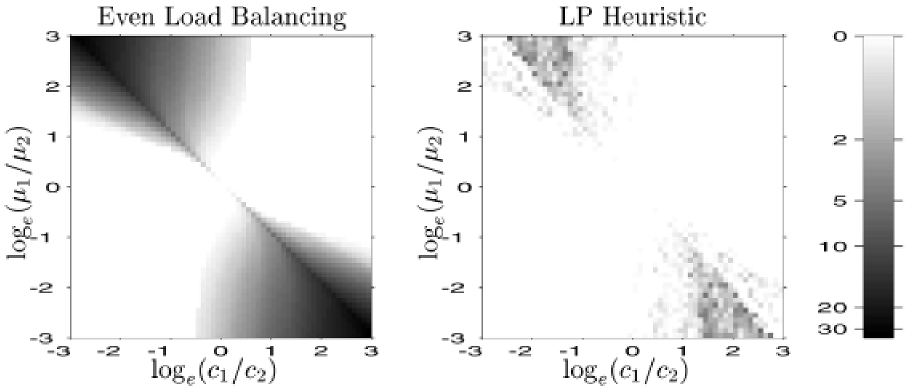

Fig. 2. Maximum suboptimality (percentage of optimum)

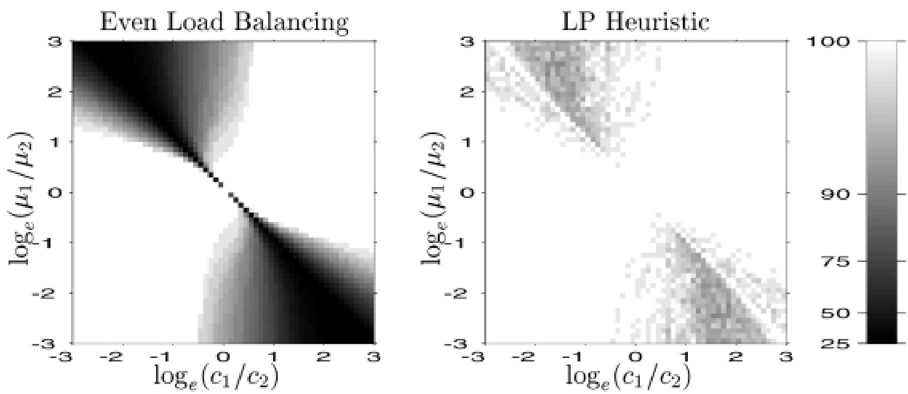

Fig. 3. Percentage of solutions optimal

where ϕ is a convex function of $\boldsymbol{\lambda}$ and where ψ is NE convex, but not convex in general. Hence we might expect the degree of non-convexity of Z^{OPT} to be related to the absolute size of $c_1\mu_1 - c_2\mu_2$ and to be at its most pronounced when $\log_e(c_2/c_1) \cong -\log_e(\mu_2/\mu_1)$.

(c) By theorem 4, $\log_e(\mu_2/\mu_1) = 0$ is a convex case for which the even load distribution will be optimal.

The results presented in Figs 2 and 3 are wholly consistent with (a)–(c). The even load distribution heuristic is optimal when $\log_e(\mu_2/\mu_1) = 0$ and has its weakest performance around the line $\log_e(c_2/c_1) = -\log_e(\mu_2/\mu_1)$. The LP-based heuristic offers a significant improvement when non-convexity is a serious issue and achieves a high level of performance almost uniformly.

In addition to the special role of even load distributions for identical stations, simple algorithms for load balancing are available when $|G| = 1$ and the $Z_m^{\mathrm{OPT}}(\cdot)$ are fully convex, but where stations are not identical. The latter raises many important modelling possibilities, including those in which the dedicated traffic has a different stochastic character at different stations and also where the processing time distributions of generic jobs are station dependent. The algorithms concerned are all based on procedures which match gradients and are variants of those proposed by Tantawi and Towsley (1984, 1985). We omit the details.

We conclude by supposing that we have such an algorithm for a network in which corollary 2 applies at each station — namely there is a single generic class and the minimized cost is increasing and convex. Throughout this discussion we shall require that all stations have no generic feed-back and give D-jobs pre-emptive priority over G-jobs. With this set-up, balancing the generic load can have no effect on the total costs incurred by dedicated jobs across the network. In our algorithm for the $|G| = 1$ case, $\nu_m(\lambda)$ is the optimal generic load at station m when λ is the total generic load for the network. See Dacre and Glazebrook (1999) for a proof of the following lemma.

Lemma 1. There is a solution to the above simple load balancing problem with $|G| = 1$ for which ν_m is increasing for each m.

We shall show how to use such an algorithm for the $|G| = 1$ case to develop an algorithm for the more general case considered in theorem 4, where the $|G|$ generic job classes have processing requirements which are IID at each station (but not necessarily identically distributed at different stations). We shall require that we have holding cost rates c_g, $g \in G$, which apply across the network. It is straightforward to establish that the optimal scheduling of generic jobs at each station is according to priorities determined by the c_g, $g \in G$, with the largest c_g having the highest priority. Renumber the generic classes such that

$$c_{|G|} \geqslant c_{|G|-1} \geqslant \ldots \geqslant c_1.$$

Recall that the total generic loads for the network are λ_g, $g \in G$. Theorem 6 describes an optimal load balancing regime for this situation.

Theorem 6. There is a solution to the above load balancing problem for which the optimal class g load at station m is

$$\nu_m\left(\sum_{j=g}^{|G|} \lambda_j\right) - \nu_m\left(\sum_{j=g+1}^{|G|} \lambda_j\right)$$

for all g and m where ν_m is as in lemma 1.

Proof. Let $\pi_{mg}(\boldsymbol{\lambda})$ be the class g load allocated to station m by a general solution π to our load balancing problem. If we write $Z(\pi, \boldsymbol{\lambda})$ for the total network cost for the generic jobs under this solution then it is not difficult to see that we have the decomposition

$$Z(\pi, \boldsymbol{\lambda}) = \sum_{g=1}^{|G|} Z_g(\pi, \boldsymbol{\lambda}). \tag{48}$$

In equation (48), $Z_g(\pi, \boldsymbol{\lambda})$ is the generic cost associated with an equivalent $|G| = 1$ network in which the station m load is $\Sigma_{j=g}^{|G|} \pi_{mj}(\boldsymbol{\lambda})$ and the common holding cost rate is $c_g - c_{g-1}$. We take $c_0 = 0$. But

$$\sum_{m=1}^{M} \sum_{j=g}^{|G|} \pi_{mj}(\boldsymbol{\lambda}) = \sum_{j=g}^{|G|} \lambda_j$$

is the total load for this network, and so by lemma 1 it is optimized by allocating generic load $\nu_m(\Sigma_{j=g}^{|G|} \lambda_j)$ to each station m. However,

$$\nu_m\left(\sum_{j=g}^{|G|} \lambda_j\right) = \sum_{j=g}^{|G|} \nu_{mj}(\boldsymbol{\lambda}) \qquad \text{for all } g \text{ and } m$$

where $\nu \equiv \{\nu_{mj}\}$ is the load balancing solution proposed in the theorem. Note that lemma 1 guarantees the admissibility of ν. From this we conclude that

$$Z_g(\pi, \boldsymbol{\lambda}) \geqslant Z_g(\nu, \boldsymbol{\lambda}) \qquad \text{for all } g$$

and so, from equation (48),

$$Z(\pi, \boldsymbol{\lambda}) \geqslant Z(\nu, \boldsymbol{\lambda}),$$

as required. □

4.3. Extensions

Dacre and Glazebrook (1999) have described general conditions which guarantee that a GCL system has an increasing and supermodular value function Z^{OPT}. Systems are described which meet the requirements and these include the Klimov network model of this section.

Most of the above discussion via the Klimov network model supposes that the D-customers and the G-customers at a station are dealt with on the same basis through a linear objective involving all job classes. Hence, prioritizing between these two customer types (and the natural proposal is to give dedicated customers a higher priority) is via an appropriate choice of the c_i, $i \in D \cup G$. Another obvious approach is to *impose* the requirement that D-customers must always be given priority over G-customers as is done in the concluding discussion leading to theorem 6. The generic customers then have the status of 'background' jobs which are allowed access to service capacity which is surplus to the primary goal of serving the D-customers. This proposal can easily be accommodated through Garbe and Glazebrook's (1998b) achievable region account of stochastic scheduling with imposed priorities.

Another way of asserting the primacy of the D-customers at each station is to impose delay constraints of the form $x_d^u \leqslant t_d$, $d \in D$. Among the controls which meet the delay constraints the goal would be to choose one to minimize $\Sigma_{g \in G} c_g x_g^u$. This is the approach of Ross and Yao (1991) who took as their station model a multiclass $M/G/1$-queue with priorities imposed non-pre-emptively. In this case we have convexity of the optimal returns for the case of a single generic class.

5. Threshold policies for intensity control

A fundamental question which can be asked of any queuing system concerns how much work we need to hold in the system to achieve a given level of throughput (i.e. the rate of job flow through the system). The intuition is that letting the work in process (WIP) grow beyond a certain level will do little to increase throughput. However, achieving a given throughput can only be done at the expense of sufficiently large WIP. A class of policies used frequently in practice is the class of *threshold policies* which control the system by setting a WIP cap. When the WIP reaches this cap the arrivals process is shut off. The following basic questions arise.

What is the minimum WIP level required to attain a target throughput level? When are threshold policies optimal for maximizing a linear (or, more generally, a convex) throughput–WIP objective?

There is a large body of research establishing the optimality of threshold policies in a variety of queuing intensity control models. The main approach is based on demonstrating certain structural properties (submodularity) of the optimal value function, through the analysis of the DP optimality equations. See Glasserman and Yao (1994). In a related approach, threshold optimality is established by using LP duality arguments. The LP formulations concerned are closely related to DP and typically have a large or infinite number of variables. See Yao and Schechner (1989). As explained in Section 1, the achievable region approach seeks to construct formulations in a *projected* variable space of *natural* performance measures. This typically involves a dramatic reduction in dimensionality, with resulting computational advantages for dealing with, for example, side-constraints. A general achievable region framework for threshold optimality, along these lines, is given in Niño-Mora (1998). This section contains an introduction to the key ideas based on an application to a queuing intensity control model due to Chen and Yao (1990). At the end of the section we give an indication of how the ideas generalize.

The model is a queuing system which consists of a facility servicing a single customer class. $N(t)$ denotes the number of customers in the system at time $t \geqslant 0$. We control the process $\{N(t),\ t \geqslant 0\}$ by means of a policy which sets the current *stochastic intensities* (or rates) $\lambda(t)$ and $\mu(t)$ of the arrival and departure processes respectively. The sequences $\{\bar{\lambda}_k,\ k = 0, 1, \ldots\}$ and $\{\bar{\mu}_k,\ k = 0, 1, \ldots\}$ of input and output capacity limits impose bounds on the arrival and departure intensities when k customers are in the system. A policy will be *admissible* if it is non-anticipative (i.e. it is adapted to the system's history), stable (i.e. the process $\{N(t),\ t \geqslant 0\}$ is ergodic) and satisfies the input and output capacity constraints, expressed as

$$N(t) = k \Longrightarrow \lambda(t) \leqslant \bar{\lambda}_k, \quad \mu(t) \leqslant \bar{\mu}_k, \qquad t \geqslant 0,\ k = 0, 1, 2, \ldots.$$

We denote by $\mathcal{U}$ the class of admissible policies. Of special interest is the class of *threshold policies*: for each integer $b \geqslant 0$, the *b-threshold policy* sets the input intensity at full capacity if $N(t) < b$, and to 0 otherwise. The output intensity is always set at full capacity.

The achievable region approach requires us to develop a notion of performance, which here must measure both throughput (μ^u for policy u) and WIP (N^u). We consider a time-average criterion and define

$$\mu^u = \lim_{T\to\infty} \left[\frac{1}{T} \int_0^T E_u\{\mu(t)\}\, \mathrm{d}t \right] \tag{49}$$

and

$$N^u = \lim_{T\to\infty} \left[\frac{1}{T} \int_0^T E_u\{N(t)\}\, \mathrm{d}t \right]. \tag{50}$$

We focus primarily on the following economic structure: a unit reward is received at each service completion time. In addition, each customer in the system (whether waiting or in service) incurs holding costs at a rate $c > 0$ per unit time. Our goal is to choose an admissible control to maximize the long-term net rate of return, i.e.

$$Z^{\mathrm{OPT}}(c) = \sup_{u\in\mathcal{U}}\{Z^u(c)\} = \sup_{u\in\mathcal{U}}(\mu^u - cN^u). \tag{51}$$

To make progress we need the following plausible assumptions about the input and output capacity limits. In this assumption, the terms increasing and decreasing are used in the non-strict sense:

(a) the sequence $\{\bar{\lambda}_k, k \geqslant 0\}$ of input capacity limits is decreasing;
(b) the sequence $\{\bar{\mu}_k, k \geqslant 0\}$ of output capacity limits is increasing concave, with $\bar{\mu}_{k+1} - \bar{\mu}_k \to 0, k \to \infty$.

Consider now this system evolving under a b-threshold policy, defined above. We denote the associated performance measures μ^b and N^b, and the corresponding objective $Z^b(c) = \mu^b - cN^b$, $b \geqslant 0$. We also write c^b for the *critical cost parameter*, given by

$$c^b = (\mu^b - \mu^{b-1})/(N^b - N^{b-1}), \qquad b \geqslant 1, \tag{52}$$

with $c^0 = 0$. Under a b-threshold policy the system evolves as a birth–death process on states $0, \ldots, b$ with state-dependent birth intensities $\bar{\lambda}_i$, $0 \leqslant i \leqslant b-1$ (and 0 otherwise), and death intensities $\bar{\mu}_i$, $1 \leqslant i \leqslant b$. The stationary distribution of this process is well known to be given by

$$\pi_i^b = K_b \prod_{j=0}^{i-1} \bar{\lambda}_j / \bar{\mu}_{j+1}, \qquad 0 \leqslant i \leqslant b, \tag{53}$$

where an empty product is 1 and K_b is the required normalizing constant. We have

$$\mu^b = \sum_{i=1}^{b} \bar{\mu}_i \pi_i^b, \qquad N^b = \sum_{i=1}^{b} i \pi_i^b, \qquad b \geqslant 1. \tag{54}$$

An expression for the critical cost parameter c^b is easily recovered from equations (53) and (54).

It is straightforward to demonstrate that the following properties of the quantities introduced above flow from the assumptions. See Niño-Mora (1998) for details.

Lemma 2.

(a) The sequences $\{\mu^b, b \geqslant 0\}$ and $\{N^b, b \geqslant 0\}$ are both (strictly) increasing.
(b) The sequence $\{c^b, b \geqslant 1\}$ is positive and (strictly) decreasing with limit 0.
(c) $\mu^u - c^b N^u \leqslant Z^b(c^b), \quad u \in \mathcal{U}, \quad b \geqslant 1$.

Note that lemma 2, part (c), is an assertion of the optimality of the b-threshold policy for the critical cost parameter c^b, $b \geqslant 1$. The achievable region analysis of the stochastic optimization problem (51) for any $c > 0$ now flows naturally. Many of the issues raised in the introductory paragraph to this section are resolved as a by-product of the analysis.

We introduce the *performance space*

$$X = \{(\mu^u, N^u), u \in \mathcal{U}\}.$$

Following lemma 2, a natural candidate for X is the *threshold polygon* P given by

$$P = \{\mathbf{x} \in (\Re^+)^2; x_1 - c^b x_2 \leqslant Z^b(c^b), \text{ for } b \geqslant 1\} \tag{55}$$

which is depicted in Fig. 4. It is easy to show that the extreme points on the lower boundary of P are (μ^b, N^b), $b \geqslant 1$, namely the performances of the b-threshold policies. The corresponding LP of interest is given by

$$Z^{\mathrm{LP}}(c) = \max_{\mathbf{x} \in P}(x_1 - cx_2). \tag{56}$$

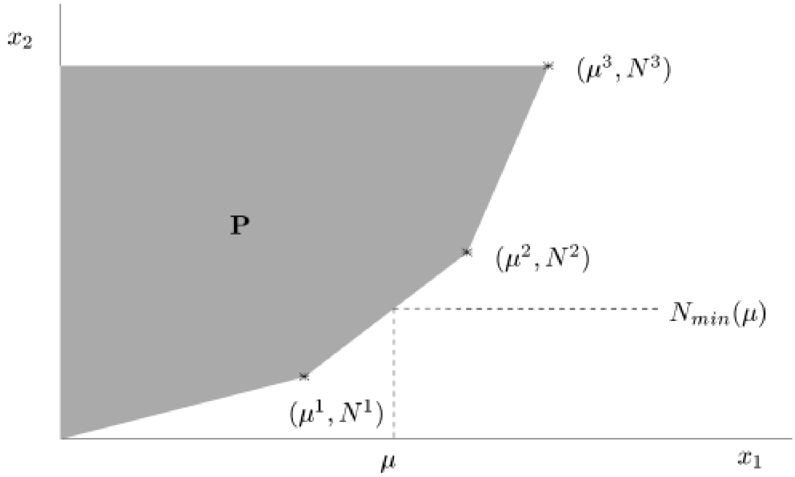

Fig. 4. Threshold polygon *P*

In our main result we shall require the *critical threshold function* $b^*(\cdot)$ given by

$$b^*(c) = \min\{b \geqslant 0;\ c^{b+1} \leqslant c\}.$$

Theorem 7 (threshold optimality via the achievable region).

(a) $Z^{\mathrm{LP}}(c) = \mu^{b^*(c)} - cN^{b^*(c)}$, $c > 0$.
(b) $X \subseteq P$.
(c) $Z^{\mathrm{LP}}(c) = Z^{\mathrm{OPT}}(c)$, $c > 0$.
(d) The stochastic optimization problem (51) is solved by the $b^*(c)$-threshold policy, $c > 0$.
(e) $X = \bar{P}$, the closure of P.

Outline proof. Part (a) follows by considering the dual LP of equation (56) through a standard complementary slackness argument which makes use of the properties described in lemma 2, parts (a) and (b).

Part (b) is an immediate consequence of lemma 2, part (c).

For part (c), it follows from part (b) that $Z^{\mathrm{OPT}}(c) \leqslant Z^{\mathrm{LP}}(c)$. However, from part (a), $Z^{\mathrm{LP}}(c)$ is achieved by the performance $(\mu^{b^*(c)}, N^{b^*(c)})$ of the $b^*(c)$-threshold policy. This yields $Z^{\mathrm{LP}}(c) \leqslant Z^{\mathrm{OPT}}(c)$ and part (c) follows.

Part (d) is an immediate consequence of parts (a) and (c).

For part (e), plainly, from part (b) we have that $X \subseteq \bar{P}$. To secure the reverse inclusion, the reader is referred to Fig. 4 for assistance. Observe that any point on the lower boundary of P is the performance of a policy which randomizes between (at most) two threshold policies. Hence the lower boundary of P is contained in X. Note also that all points $(0, N)$ are in X where N is a non-negative integer. To see this, consider a policy which guarantees that the system enters the state in which N customers are present in finite time and which then freezes the system by closing down both the input and the output. Plainly there is such a policy and its performance is $(0, N)$.

By appealing further to randomizations, we infer that the convex hull of the lower boundary of P together with $\{(0, N), N \geqslant 0\}$ is contained in X. We deduce that $\bar{P} \subseteq X$ and part (e) follows. □

We now broach the issue raised above of the minimum WIP level $N_{\min}(\mu)$ required to achieve a target throughput level μ. From theorem 7 we can write

$$\begin{aligned} N_{\min}(\mu) &= \min\{N^u;\ \mu^u = \mu,\ u \in \mathcal{U}\} \\ &= \min\{N;\ (\mu, N) \in X\} \\ &= \min\{N;\ (\mu, N) \in \bar{P}\}. \end{aligned} \tag{57}$$

The minimization in equation (57) is achieved on the lower boundary of P. See Fig. 4. Corollary 4 follows easily. We write $\mu^\infty = \lim_{b\to\infty}(\mu^b)$.

Corollary 4. $N_{\min}(\mu)$ is a piecewise linear function of μ over the range $[0, \mu^\infty)$ given by

$$N_{\min}(\mu) = N^b + \frac{1}{c^{b+1}}(\mu - \mu^b), \qquad \mu^b \leqslant \mu \leqslant \mu^{b+1}, \quad b \geqslant 0.$$

We note that the above analysis allows us to consider a *convex optimization* problem which generalizes equation (51) as follows. Let f and g be increasing, positive-valued functions with continuous derivatives, with f concave and g convex. Now the reward corresponding to policy u is given by $f(\mu^u) - g(N^u)$, yielding the stochastic optimization problem

$$Z^{\text{OPT}}_{\text{convex}} = \sup_{u\in\mathcal{U}}\{f(\mu^u) - g(N^u)\}. \tag{58}$$

Following theorem 7 we can reformulate problem (58) as the convex optimization problem

$$Z^{\text{OPT}}_{\text{convex}} = \max_{x\in P}\{f(x_1) - g(x_2)\} \tag{59}$$

with P given in equation (55). However, by standard results in convex optimization (see, for example, chapter 2 in Bertsekas (1995)) we know that a necessary and sufficient condition for the achievable throughput–WIP pair (μ^*, N^*) to solve problem (59) is that it solves the LP

$$Z^{\text{LP}}\left\{\frac{g'(N^*)}{f'(\mu^*)}\right\} = \max_{x\in P}\left\{x_1 - \frac{g'(N^*)}{f'(\mu^*)}x_2\right\}. \tag{60}$$

Theorem 8 can be shown to follow easily from theorem 7. In this result, the threshold level b^* is given by

$$b^* = \min\left\{b \geqslant 0;\ c^{b+1} \leqslant \frac{g'(N^{b+1})}{f'(\mu^{b+1})}\right\}.$$

The quantity α^* then solves the convex optimization problem

$$\max_{\alpha\in[0,\,1]}[f\{\alpha\mu^{b^*} + (1-\alpha)\mu^{b^*+1}\} - g\{\alpha N^{b^*} + (1-\alpha)N^{b^*+1}\}].$$

Theorem 8 (optimal intensity with convex objective). The randomized policy that selects, at time 0, the b^*-threshold policy with probability α^* and the (b^*+1)-threshold policy with probability $1-\alpha^*$ is optimal for control problem (58).

Consider now a general stochastic system with performance summarized by a throughput–WIP pair (μ^u, N^u). The stochastic optimization problem of interest is equation (51) and the general system continues to be furnished with a class of threshold policies whose associated performances are (μ^b, L^b), $b \geqslant 0$. Niño-Mora (1998) describes what needs to be true in general of the set $\{(\mu^b, N^b), b \geqslant 0\}$ for the achievable region to be a *threshold polygon* (as in the above example) whose vertices are the performances of threshold policies. When these conditions are met, threshold policies will be optimal for linear objectives and the minimum WIP level $N_{\min}(\mu)$ will be piecewise linear, as in corollary 4.

6. Comments and plans for future work

We make no claim that the achievable region approach will always be the preferred option in tackling stochastic optimization problems. It certainly appears to be a more than useful addition to the toolkit. Direct comparisons with existing approaches are not straightforward. In DP, value iteration and policy improvement algorithms are (virtually) routinely available and there is nothing equivalent in the achievable region approach where plenty of creative thinking may be involved in a succesful application of the ideas. That said, these are early days and already certain approaches to the development of constraints on performance variables seem to be especially productive. In multiclass queuing systems such as those of Sections 3 and 4 where the performances are given by expectations taken with respect to a stationary distribution, work decomposition laws such as equation (28) have played a central role. Such laws have been developed by appeal to principles of flow conservation in the systems concerned and also by the application of methods based on potential functions, as here. See Glazebrook and Niño-Mora (1997).

All the examples discussed in this paper involve objectives expressed in terms of time averages, but the approach is certainly not limited in its usefulness to such problems. Branching bandit problems (including multiarmed bandits) with discounted costs have been analysed by using these methods. There is no reason in principle why the approach could not be applied to finite horizon problems.

Current plans for further development of the achievable region approach by the authors and co-workers include work in the following three major areas.

6.1. Primal–dual approach

As mentioned at the end of Section 2, the methodology underlying the performance guarantee in corollary 1 is derived from the primal–dual structure of LPs. The method works by constructing both a heuristic solution to an appropriately defined (primal) LP related to the stochastic optimization problem of interest and a feasible solution to the dual of a relaxation of it. Our goal is to establish this approach as the central methodology in the analysis of heuristic policies for the control of stochastic systems within achievable region methodology both by extending its application to approximately GCL systems (like those discussed in Section 3) and by introducing it as an analytical tool in new contexts, including the intensity control problems of Section 5.

6.2. Load balancing

There is huge scope for further development of the work in Section 4. We shall mention just two directions for such work: firstly, the delay-constrained problem of Ross and Yao (1991) mentioned at the conclusion of Section 4 is both compelling from the perspective of applications, but also a formidable technical challenge when placed in the context of GCL systems. Secondly, in more complex systems than those discussed above for which the model for each station only approximately satisfies GCLs then functions $\tilde{Z}_m$ approximating the optimal costs Z_m^{OPT} will have the kind of convexity properties discussed in Section 4. A natural load balancing heuristic can be obtained by replacing Z_m^{OPT} by $\tilde{Z}_m$ in problem (41). Further work will include the development of performance guarantees for such heuristic approaches.

6.3. Extension of the approach to new areas

Strict priority policies for the service of customers in a queuing network may be unattractive because of the heavy penalties that they impose on low priority jobs or customers. The latter suffer not only large queues and response times but, perhaps more significantly, large variances in these quantities. Natural formulations to ameliorate this would seek policies to minimize the usual time-average linear holding cost rate subject to constraints on the variance or to incorporate quadratic terms in the objective. We have begun work in this challenging area and believe that the achievable region approach has an important role to play. Achievable region methodology will also be introduced as an analytical tool for developments of the models discussed in Section 5 to accommodate scheduling of the WIP in addition to the control of the arrivals process.

Acknowledgements

We would like to express our appreciation to the Engineering and Physical Sciences Research Council for supporting the work of the first author by means of a research studentship and for supporting the work of the second author through the award of grants GR/K03043 and GR/M09308. The work of the third author has been supported by European Community Marie Curie Research Fellowship ERBFMBICT961480, and by a Research Initiation Award at Universitat Pompeu Fabra. We would also like to thank colleagues in the Department of Statistics, Newcastle University, for their constructive comments on earlier drafts of the paper and the referees for their helpful suggestions.

References

Atkins, D. and Chen, H. (1995) Performance evaluation of scheduling control of queueing networks: fluid model heuristics. *Queueing Syst*., **21**, 391–413.

Bertsekas, D. P. (1995) *Nonlinear Programming*. Belmont: Athena Scientific.

Bertsimas, D. (1995) The achievable region method in the optimal control of queueing systems—formulations, bounds and policies. *Queueing Syst*., **21**, 337–389.

Bertsimas, D. and Niño-Mora, J. (1996) Conservation laws, extended polymatroids and multi-armed bandit problems: a polyhedral approach to indexable systems. *Math. Oper. Res*., **21**, 257–306.

Bertsimas, D., Paschalidis, I. C. and Tsitsiklis, J. N. (1995) Branching bandits and Klimov's problem: achievable region and side constraints. *IEEE Trans. Autom. Control*, **40**, 2063–2075.

Chang, C. S. and Yao, D. D. (1993) Rearrangement, majorization and stochastic scheduling. *Math. Oper. Res*., **18**, 658–684.

Chen, H. and Yao, D. D. (1990) Optimal intensity control of a queueing system with state-dependent capacity limit. *IEEE Trans. Autom. Control*, **35**, 459–464.

Coffman, E. and Mitrani, I. (1980) A characterization of waiting time performance realizable by single server queues. *Oper. Res*., **28**, 810–821.

Dacre, M. J. and Glazebrook, K. D. (1999) An approach to load balancing in distributed systems. To be published.

Federgruen, A. and Groenevelt, H. (1988) Characterisation and optimisation of achievable performance in queueing systems. *Oper. Res*., **36**, 733–741.

Garbe, R. and Glazebrook, K. D. (1998a) Submodular returns and greedy heuristics for queueing scheduling problems. *Oper. Res*., **46**, 336–346.

——— (1998b) Stochastic scheduling with priority classes. *Math. Oper. Res*., **23**, 119–144.

Gelenbe, E. and Mitrani, I. (1980) *Analysis and Synthesis of Computer Systems*. London: Academic Press.

Gelenbe, E. and Pekergin, F. (1993) Load balancing pragmatics. *Technical Report*. Université René Descartes.

Gittins, J. C. (1979) Bandit processes and dynamic allocation indices (with discussion). *J. R. Statist. Soc.* B, **41**, 148–177.
——— (1989) *Multi-armed Bandit Allocation Indices*. New York: Wiley.
Glasserman, P. and Yao, D. D. (1994) Monotone optimal control of permutable GSMPs. *Math. Oper. Res.*, **19**, 449–476.
Glazebrook, K. D. (1982) On a sufficient condition for superprocesses due to Whittle. *J. Appl. Probab.*, **19**, 99–110.
Glazebrook, K. D. and Garbe, R. (1999) Almost optimal policies for stochastic systems which almost satisfy conservation laws. *Ann. Oper. Res.*, to be published.
Glazebrook, K. D. and Niño-Mora, J. (1997) Scheduling multi-class queueing networks on parallel servers: approximate and heavy traffic optimality of Klimov's rule. *Lect. Notes Comput. Sci.*, **1284**, 232–245.
Harrison, J. M. (1996) The BIGSTEP approach to flow management in stochastic processing networks. In *Stochastic Networks: Theory and Applications* (eds F. Kelly, S. Zachary and I. Ziedens), pp. 57–90. Oxford: Oxford University Press.
Harrison, J. M. and Nguyen, V. (1993) Brownian models of multiclass queueing networks: current status and open problems. *Queueing Syst. Theory Applic.*, **13**, 5–40.
Harrison, J. M. and Wein, L. (1989) Scheduling networks of queues: heavy traffic analysis of a simple open network. *Queueing Syst.*, **5**, 265–280.
Klimov, G. P. (1974) Time sharing systems I. *Theory Probab. Applic.*, **19**, 532–551.
Liu, Z. and Towsley, D. (1994) Optimality of the round-robin routing policy. *J. Appl. Probab.*, **31**, 466–478.
Maglaras, C. (1997) Design of dynamic control policies for stochastic processing networks via fluid models. In *Proc. CDC97*. New York: Institute of Electrical and Electronic Engineers. To be published.
Manne, A. (1960) Linear programming and sequential decisions. *Mangmnt Sci.*, **6**, 259–267.
Niño-Mora, J. (1998) On the throughput-WIP trade-off in queueing systems, diminishing returns and the threshold property: a linear programming approach. *Economics Working Paper 276*. Universitat Pompeu Fabra, Barcelona.
Ross, K. W. and Yao, D. D. (1991) Optimal load balancing and scheduling in a distributed computer system. *J. Ass. Comput. Mach.*, **38**, 676–690.
Shanthikumar, J. G. and Yao, D. D. (1992) Multi-class queueing systems: polymatroidal structure and optimal scheduling control. *Oper. Res.*, **40**, 293–299.
Tantawi, A. N. and Towsley, D. (1984) A general model for optimal static load balancing in star network configurations. In *Performance '84*, pp. 277–291.
——— (1985) Optimal static load balancing in distributed computer systems. *J. Ass. Comput. Mach.*, **32**, 445–465.
Walrand, J. (1988) *An Introduction to Queueing Networks*. Englewood Cliffs: Prentice Hall.
Weber, R. R. (1978) On the optimal assignment of customers to parallel queues. *J. Appl. Probab.*, **15**, 406–413.
——— (1992) On the Gittins index for multi-armed bandits. *Ann. Appl. Probab.*, **2**, 1024–1033.
Weber, R. R. and Weiss, G. (1990) On an index policy for restless bandits. *J. Appl. Probab.*, **27**, 637–648.
Weiss, G. (1988) Branching bandit processes. *Probab. Engng Inform. Sci.*, **2**, 269–278.
——— (1992) Turnpike optimality of Smith's rule in parallel machine stochastic scheduling. *Math. Oper. Res.*, **17**, 255–270.
——— (1995) On almost optimal priority rules for preemptive scheduling of stochastic jobs on parallel machines. *Adv. Appl. Probab.*, **27**, 821–839.
Whittle, P. (1980) Multi-armed bandits and the Gittins index. *J. R. Statist. Soc.* B, **42**, 143–149.
——— (1988) Restless bandits; activity allocation in a changing world. *J. Appl. Probab.* A, **25**, 287–298.
Yao, D. D. and Schechner, Z. (1989) Decentralized control of service rates in a closed Jackson network. *IEEE Trans. Autom. Control*, **34**, 236–240.

Discussion on the paper by Dacre, Glazebrook and Niño-Mora

R. J. Gibbens (*University of Cambridge*)

The authors are to be congratulated on a stimulating and enjoyable account of the admissible region approach to stochastic control problems. The family of problems tackled, i.e. multiclass server systems, presents a significant challenge for analysis. Moreover, such systems are of increasing importance in the practical modelling world, especially that of communication systems.

In today's communication networks there is a pressing need to integrate varying traffic types such as voice, data and video to share common network resources. Yet these traffic types have widely differing statistical characteristics and correspond to services with widely differing requirements for quality or performance as measured by individual packet loss or delay. Multiclass models naturally emerge from these engineering designs and any insights given by a modelling analysis can produce significant benefits to the overall system. Often these benefits will be in the form of simpler and hence cheaper mechanisms as well, giving higher performance.

One topic that I should like to discuss is what if the performance space X is difficult to identify (step

(a) of the *achievable region approach*)? Consider an example from an investigation of network routing (for further details see Gibbens and Kelly (1995)). In this example, calls of class i arrive at the network at rate ν_i ($i \in I$). The resources in the network consist of links and let link i have a capacity of C_i circuits. A call of class i may be routed directly on link i or alternatively on a route $r \in R(i)$ (r is a subset of I and write $R = \cup_{i \in I} R(i)$). An arriving call that is accepted holds a circuit on each of the links for the holding period of the call. A call may only be assigned to a route r with at least one free circuit from each link $j \in r$. The call may also be rejected (or blocked), and it must be rejected if there are no routes with spare capacity. The holding period of a call is exponentially distributed with unit mean and is unaffected by the route used to carry the call. In networks with plenty of routing choices then X may be difficult to identify. In such cases it is possible that good bounds on X may provide some insight into the structure of optimal policies.

Here $X \subseteq P$ where the polyhedron P is the set of vectors $(f_i,\ i \in I)$ satisfying

$$f_i = d_i + \sum_{r \in R(i)} e_r \leqslant \nu_i \qquad i \in I,$$

$$\sum_{r \ni i} e_r \leqslant M_i(d_i) \qquad i \in I$$

for some $d_i \geqslant 0$, $i \in I$ and $e_r \geqslant 0$, $r \in R$.

The function $M_i(d_i)$ is the piecewise linear concave function which gives the maximal mean acceptance rate of alternatively routed calls on link i, subject to the requirement that the mean acceptance rate of directly routed calls is at least d_i. This is given by the solution of a one-dimensional Markov decision problem and is of simple threshold type. Thus the expected number of calls carried per unit time is bounded above by the solution to

$$\max\left(\sum_{i \in I} f_i\right)$$

subject to

$$(f_i,\ i \in I) \in P.$$

How close is the bound? Hunt and Laws (1993) have established that for a symmetric network as the number of nodes n increases then the proportion of calls lost under an optimal dynamic routing scheme approaches the loss given by the bound. Moreover, the form of the optimal scheme is of a simple threshold type (known as trunk reservation), corresponding to the solution of the one-dimensional MDP giving the functions $M(d)$. For asymmetric networks the bound can be calculated and has been found to be approached by simple decentralized dynamic routing schemes (see Gibbens and Reichl (1995)).

In conclusion, it is my pleasure and privilege to propose a vote of thanks to the authors.

P. Whittle (*University of Cambridge*)

It is a pleasure to have heard such a stimulating paper, and I congratulate the authors.

Many optimal service policies show a particular simplification under the steady state criterion of average optimality, and it is just in this case that the achievable region–linear program (LP) approach seems to be particularly natural.

As the authors make plain, the essential step in the approach is the determination of the achievable region X. This seems to be surprisingly difficult, and the authors' generalized conservation law characterization on pages 752 and 758 shows great ingenuity — all the more, for this characterization is far from evident *a priori*.

I cannot escape the feeling that the determination of X should be more evident in these largely linear cases, and also that behind every model for which an index policy is almost optimal lies a modified model for which it is exactly optimal. In my treatment of the restless bandit problem (Whittle, 1988), to which the authors refer, a very transparent LP formulation emerged once the problem had been relaxed in an appropriate way. It turns out that, with such a relaxation, the problems that the authors consider on pages 752 and 756 are just special cases of restless bandits, and that the Klimov indices can be deduced in literally a couple of lines.

In what one might call the *service relaxation* the only constraint which is distinguished explicitly is the restriction to M servers. This constraint is relaxed simply by costing service at a fixed rate ν per unit time. The values of ν which are critical if a customer of class j is to be served or not define the generalized Gittins index ν_j for such a customer. The charge rate ν is the Lagrange multiplier associated with the prescription of the *average* number of servers which are active in equilibrium. The actual number which are active will vary randomly in time, but the constraint is satisfied essentially and the index policy is optimal in an asymptotic sense conjectured by Whittle (1988) and established by Weber and Weiss (1990).

So, for example, consider the Klimov model treated on pages 756–761 of the paper: a multiclass queuing model in which customers may depart on completion of service or may be recycled with a possible change of class. Let us phrase the model in continuous time with a finite pool of customers, who are *engaged* or *disengaged* according to whether they require service or not. Denote the engagement (arrival) rate, disengagement (service) rate and cost rate for a prescribed customer of class j by λ_j, μ_j and c_j respectively. The probability that an engaged customer of class j is recycled as one of class k on completion of service is denoted p_{jk}, and the probability that on completion of service he is restored to the pool as a customer of class k is denoted q_{jk}. If the matrices with these elements are denoted P and Q then $P+Q$ is a stochastic matrix.

The relaxed process is still dynamic, and the optimality condition will be a dynamic programming equation, but the effect of the relaxation is to decouple the customers. The dynamic programming equation becomes simply

$$\gamma_j = \lambda_j(f_j - g_j) = \min\left\{c_j,\ c_j + \nu - \mu_j f_j + \mu_j \sum_k (p_{jk} f_k + q_{jk} g_k)\right\} \tag{61}$$

where γ_j is the minimal expected cost incurred per unit time by a customer of class j, f_j is the marginal increase in cost if a customer of class j is added to the queue and g_j is the marginal increase in cost if a customer of class j is added to the pool. The first and second equalities of equation (61) constitute the dynamic programming equation in the cases where a customer of class j is disengaged or engaged respectively.

If there is no recycling or change of type, so that $P=0$ and $Q=I$, then it follows immediately from equation (61) that the index values are $\nu_j = \mu_j c_j/\lambda_j$, just the priority index for the well-known average optimal policy. The use of this index is optimal in the case $M=1$, it is close to optimal for larger M (in a sense specified by the authors and others) and it is optimal in the doubly average sense (per unit of time and per individual in the pool) if M and the pool are allowed to grow in proportion (Weber and Weiss, 1990).

In other cases equation (61) for the f_j, g_j and γ_j is best solved by starting with ν large and then decreasing it to locate the break points (and so the index values ν_j). For ν sufficiently large the solution will be simply $\gamma_j = \lambda_j(f_j - g_j) = c_j$, corresponding to the situation where nobody is served and the costs of an entire community in the queue are accepted. (The National Health Service springs to mind.) As ν is decreased then the two bracketed quantities will become equal for some j at some point, the j-value indicating the customer class of highest priority and the ν-value the corresponding index. Actually, we do not need to calculate the index; the order of j-values at which the break points occur as ν decreases gives the priority order of the customer classes.

In many cases changes of 'class' correspond simply to stages of processing for members of a given true class i. The average cost γ_j will then be the same for all these stages and will be a function of i alone. We could also allow class transitions while a customer is waiting in the queue, in such a way that his priority increased as he waited. This would then allow customers of initially low priority a means of escaping from the 'queuing trap' in reasonable time.

I have pleasure in seconding the motion that a vote of thanks be awarded.

The vote of thanks was passed by acclamation.

P. Ansell (*University of Newcastle*)

Consider the two-class $M/M/1$ queuing system with arrivals to queue k forming Poisson streams with parameters λ_k, $k=1, 2$, and having exponentially distributed service times with mean $1/\mu_k$, $k=1, 2$. Under these conditions, the achievable region methodology finds that the optimal policy needed to minimize a linear combination of average queue lengths, $C = c_1\, E(N_1) + c_2\, E(N_2)$, is a strict priority policy. Moreover, this will give highest priority to the customer class with the largest value of $c_k\mu_k$ (the $c\mu$-rule).

However, for many practical purposes, strict priority policies are unacceptable. This is because of the large penalties that are imposed on the low priority jobs. These cause the low priority jobs to have large expected queue lengths and response times, but, perhaps more importantly, large queue length variances.

A natural optimization problem which arises to mitigate this is

$$\text{minimize} \quad c_1\, E(N_1) + c_2\, E(N_2),$$
$$\text{var}(N_1) \leqslant \nu_1,$$
$$\text{var}(N_2) \leqslant \nu_2.$$

This leads to the study of two challenging problems, namely

(a) to obtain implementable and analytically tractable service policies which come close to minimizing the linear cost function and also satisfy the variance constraints and
(b) to obtain an exact characterization of the achievable region that contains both first and second moments (or at least a relaxation of it).

The first of these problems is addressed by describing two classes of heuristic policies: a randomized family and a threshold family. Both policy classes are characterized by a single parameter. To determine the performance measures for these heuristic policies two methods, generating functions and the power series algorithm, are employed. The power series algorithm is proving to be a powerful tool in the performance analysis of the threshold policies.

The second of the above problems is addressed by using potential functions to create sets of linear constraints and then to strengthen the characterization by incorporating semidefinite constraints. (Note that this is restricted to problems involving second-moment constraints instead of variance constraints.) This leads to a semidefinite programming problem which can be solved by using one of several packages available. See Ansell *et al.* (1999).

So far the results indicate that the family of threshold policies does well in comparison with the more commonly studied family of randomized policies. In many cases they come close to the bound obtained by the semidefinite programming formulation. Work in progress attempts to improve the tightness of the achievable region relaxation.

John Bather (*University of Sussex, Brighton*)

Statistics now deals with very large data sets and probability is applied to models of increasing complexity. Both these developments are associated with computing and information technology. This paper is welcome because it is concerned with important and difficult problems in stochastic optimization. My first reaction was that it would be too technical to follow, because I am not familiar with the background, but the authors have done their best to explain the principles of their approach and I found this very helpful.

The first example in Section 2 illustrates the possibilities and also some limitations of the approach. We are, in effect, minimizing a linear cost function on a convex set, so the solution corresponds to an extreme point. This means that the optimal policy does not depend on the state of the system, as represented by the numbers of customers $N_1(t)$ and $N_2(t)$. Presumably this is a consequence of allowing a pre-emptive switching of service, but it is not clear what happens otherwise.

The general queuing model of Section 3 shows that, under the generalized conservation law conditions, index policies are useful: either they are optimal ($M = 1$ server) or asymptotically optimal ($M \geqslant 2$ servers). It is remarkable how the validity of decision procedures based on indices has been extended since Gittins (1979). There are limitations: the indices are not usually derived from simple intuitive rules; they involve substantial computations. The exact form is also sensitive to the choice of model parameters, which makes me sceptical of properties like asymptotic optimality.

In Section 6, the authors note that there is nothing in the achievable region approach which is equivalent to the value iteration and policy improvement algorithms used in dynamic programming. Perhaps some kind of policy improvement may be feasible. This would require an investigation of relative value functions representing the advantage or disadvantage of one particular state of the system relative to another, under a given control procedure and the equilibrium distribution generated by it. Roughly speaking, local optimality with respect to a relative value function can be used to establish global optimality of the corresponding control procedure. Similarly, local suboptimality may suggest local improvements which lead to a global improvement in the policy.

Finally, I hope that this paper will stimulate as much further research as Gittins (1979) did 20 years ago.

E. J. Collins (*University of Bristol*)
The authors are to be congratulated on the clarity with which they have enunciated the achievable region approach and the success that they and their co-workers have had in using the concept of generalized conservation laws (GCLs) to unify results and methodology for various classes of stochastic optimization problems. It will be interesting to see how far the approach can be extended to other structured problems, and I wonder whether the authors can give any general principles for identifying and constructing appropriate GCLs.

However, as their intensity control example shows, the achievable region approach is not limited to cases where GCLs apply. As they say, in principle the approach can be applied to finite horizon problems or more generally to replace formulations in the space of state–action frequencies with simpler analyses in the space of natural performance variables.

For example, Collins and McNamara (1998) used an achievable region approach to analyse a class of finite horizon Markov decision processes with non-standard reward structures, which arise naturally in the context of behavioural ecology (McNamara *et al.*, 1995). Standard dynamic programming (DP) methods are not directly applicable because of the reward structure and a non-linear programming formulation in terms of state–action frequencies suffers from the problem of large dimensionality referred to above. Instead, each finite horizon policy π is associated with a system performance vector $\mathbf{x}^{\pi}$, whose components $x^{\pi}(i)$ represent the probability that the process is in state i at final time T and the overall reward is taken to be a strictly concave function ϕ of the final distribution $\mathbf{x}$. The performance space can be characterized as a convex polytope $\mathbf{X}$, each of whose vertices corresponds to the final distribution associated with some (non-stationary) Markov deterministic policy, and the optimal performance vector can be characterized implicitly by noting that $\mathbf{x}^*$ maximizes $\phi(\mathbf{x})$ if and only if

$$\nabla\phi(\mathbf{x}^*)\mathbf{x} \leqslant \nabla\phi(\mathbf{x}^*)\mathbf{x}^* \qquad \text{for all } \mathbf{x} \in \mathbf{X}.$$

The maximization problem can then be solved iteratively, without explicitly determining $\mathbf{X}$, by solving a sequence of DPs in which the finite horizon policy solving each iteration itself determines the objective function for the next iteration, and the optimal policy is in general a randomized policy of much higher dimension than the performance space. The examples presented by the authors appear to rely on the dimension of the optimal policy being no larger than that of the performance space, and I wonder whether they have any systematic way of extending their method to this more general case.

John Gittins (*University of Oxford*)
It is exciting to witness the early stages of an important new approach to a class of problems with which I am very familiar. As the authors say, it is much too early to make an informed judgment on how far the achievable region approach will take us. I should like to ask whether there seems to be any possibility of progress with two problems that I wrestled with a few years ago, before this approach was well known.

Problem 1: parallel servers
As the authors point out, Glazebrook and Garbe (1999) showed that in general an index policy performs well, particularly with heavy traffic. There are also circumstances in which an index policy is optimal, e.g. one class of customers, monotone completion rates (see Weber (1982)) and deterministic processing times, no arrival process and identical time costs (see Gittins (1989)).

Are there likely to be other circumstances for which an exact optimal solution may now be obtained?

Problem 2: index values
Does the new approach lead to new algorithms which may be used to calculate index values? I have in mind particularly the problems of multipopulation random sampling which are discussed in chapters 6 and 7 of Gittins (1989). The use of duality ideas, associated with achievable regions, is a fruitful source of effective algorithms in linear programming and combinatorial optimization, so we might hope for some progress here.

M. Zervos (*University of Newcastle*)
One of the methods to study optimal control problems is via mathematical programming on suitable spaces. Among the first results to be established by means of such an approach are Pontryagin's

maximum principle (see for example Fleming and Rishel (1975)) and the solution of the general optimal stopping problem by Snell (1952).

The paper is concerned with ramifications of such an approach to a class of stochastic optimization problems. To fix our ideas, consider the 'stationary' stochastic optimization problem

$$\min_{u \in \mathcal{U}} [E\{g(Z^u)\}], \tag{62}$$

where Z^u is the state of a controlled system in a steady state associated with admissible control u. Problem (62) is equivalent to

$$\min_{u \in \mathcal{U}} \langle g, \mu^u \rangle, \tag{63}$$

where $\mu^u = Z^u(P)$. Thus, as long as we can express the constraint that Z^u is the state of the system under control $u \in \mathcal{U}$ as a set of linear constraints on the measures μ^u (this is the challenging part), we can reformulate the original problem as a linear programming problem. Note that this is infinite dimensional. However, in the special case where $g(z) \equiv z$, the problem's objective function is expressed in terms of only the first moment of μ^u. This is the case which pertains in the paper: control u is associated with *performance* $z^u = E(Z^u)$.

With regard to their applicability, mathematical programming approaches to stochastic control have two difficulties. First, for a general choice of g, they give rise to infinite dimensional problems, which might lead to the conclusion that they are limited as far as general applications are concerned. Second, they require the additional effort of identifying a control which is associated with the solution of the resulting linear program problems, namely an optimal control. In view of major results like those mentioned in the first paragraph, these difficulties cannot be considered a criticism. Also, as far as the first is concerned, note that 'almost linear' models are related to a large class of very important applications, including those analysed in the paper. However, the analysis of large or infinite dimensional problems by means of the achievable region approach with appropriate relaxations presents a challenging issue.

As a conclusion, the method presented in the paper can be incorporated into a wider theory as an integral part. However, judging from the range of results that it has already yielded, it appears to be a most promising and important tool.

The following contributions were received in writing after the meeting.

N. Bäuerle (*University of Ulm*)

Optimal control of stochastic queuing systems has always been a challenging task. In recent years new methods have been developed for this problem — one of the most prominent is the 'achievable region approach' which has been pursued by the authors. A striking thing about all these approaches is that the stochastics of the problem do not play an important role in the analysis. Indeed, only the first moments of the relevant distributions seem to be significant. The reason may be the objective function, which is often taken to be linear.

In particular the achievable region approach is very well suited for average cost problems with linear cost functionals. The more complicated the objective function becomes, stressing the stochastic nature of the problem, the less easy it is to apply the ideas. This is also due to the 'creative thinking which may be involved in a successful application of the ideas' (p. 774). Another critical point in the philosophy of the achievable region approach is to emphasize the role of bench-mark policies such as index policies and switching policies in such a way as to characterize them as extreme policies which can be shown to be optimal for certain objectives. The application of this method has until now been restricted to models where the importance of these policies has already been known. However, many scheduling problems remain for which these known simple policies do not perform well. I wonder whether the achievable region approach can give some hints for finding new policies for these networks which are still simple and have a good performance.

However, the two big advantages of the achievable region approach in my eyes are the following. First, as demonstrated in Section 3, it is easy to make statements about how far away from the optimum simple policies like index policies are (see corollary 1). Second, further constraints on the system (like those discussed on p. 769) can easily be handled with this approach. In particular, this gives rise to new problems which have not been investigated yet. In this spirit, the achievable region approach is certainly a valuable method which should be pursued further.

John R. Birge (*University of Michigan, Ann Arbor*)
The optimization of stochastic systems is one of the most challenging yet practical problems in applied mathematics. The production, telecommunication, power, transportation, health and financial sectors all rely on decisions that seek the optimum performance in the face of significant uncertainties. This paper describes an innovative approach to these problems by identifying the set of characteristics that can be achieved by the system and by then optimizing over that region. This approach has great potential for applications in other areas but still faces challenges for applications to systems with transient or short-term characteristics.

The achievable region approach has some antecedents in other problem domains. For example, Pliska (1986) examined an optimal portfolio problem over time. The model is used to find investments at time t of θ_t in a risky portfolio and $1-\theta_t$ in a risk-free asset to maximize the expected value of the terminal wealth,

$$U(X_{\mathrm{T}}) = \int_{\omega} u\{X(\omega), \omega\}\, \mathrm{d}Q(\omega),$$

where $\omega \in \Omega$ for a probability space $(\Omega, \mathcal{F}, P)$, and where Q is the *risk neutral* or *reference* measure equivalent to P.

Pliska (1986) used the theory of complete markets established in Harrison and Pliska (1981) that the set of all *attainable wealths* X_{T} are those that satisfy $E_Q(X_{\mathrm{T}}) = 0$ (or some other initial wealth). He then considered this terminal objective and found the final distribution of the attainable wealth that maximizes this objective. After the final wealth distribution is determined, Pliska showed how to obtain a policy that achieves this distribution.

Dacre, Glazebrook and Niño-Mora show that many areas beyond finance may benefit from this approach. The problems so far, however, have considered only a single agent. Additional applications in distributed control or game theoretic models might be possible by using general equilibrium properties to define the achievable region. The results have implications for the rapidly emerging markets in communication, transportation and power.

The limitation of this approach is the ability to define the achievable region. The paper suggests that finite time horizons present no inherent difficulties but the problem in these situations is that transient characteristics often dominate and make the identification of an achievable region possible only after identifying a policy for each point in time. Such phenomena are in fact the domain of both the dynamic-programming-based approaches mentioned in the paper and stochastic programming approaches (as in, for example, Birge and Louveaux (1997)) that may be required when policies do not have a simple threshold or decision rule structure.

Esther Frostig and Gideon Weiss (*University of Haifa*)
The paper extends the achievable region approach to the derivation of approximately optimal solutions for stochastic optimization problems, by the use of approximate generalized conservation laws (GCLs). It illustrates the tremendous potential of this approach to solve hitherto intractable problems, and in particular, in Section 3, it provides the first proof of the asymptotic optimality of Klimov's policy for parallel servers.

In hindsight the achievable region approach with approximate GCL can be used to derive the performance of Smith's rule for scheduling a batch of n jobs on M machines, as given in Weiss (1992). Consider a finite batch of jobs $j \in E = \{1, 2, \ldots, n\}$ with processing times X_j and holding costs w_j per unit time. For a given scheduling policy u let $(C_1^u, C_2^u, \ldots, C_n^u)$ be the completion times of the jobs, and consider these as the performance vector of the schedule. The following approximate GCL follows from Weiss (1992) (see also Möhring *et al.* (1998)):

$$\sum_{k \in S} X_k C_k^u \geqslant b(S) \qquad S \subset E,$$

$$\sum_{k \in S} X_k C_k^u \leqslant b(S) + \Phi(S) \qquad S \subset E, \quad u\colon S \to S^c,$$

$$b(E) \leqslant \sum_{k \in E} X_k C_k^u \leqslant b(E) + \Phi(E), \tag{64}$$

where

$$b(S) = \frac{1}{2M}\left(\sum_{k\in S} X_k\right)^2 + \frac{1}{2}\sum_{k\in S} X_k^2,$$

$$\Phi(S) = \frac{(M-1)^2}{M} \max_{j\in E} (X_j^2). \tag{65}$$

Here $\Phi(S)$ is obtained as an upper bound on $\overline{D}_S^u$ which is the sample variance of the times at which the M machines fall idle at the completion of the jobs in S.

Note that these quantities are defined for a finite sample path, rather than in terms of a steady state vector of expectations. Also, because there is no feed-back of jobs, V_j^S is simply X_j.

For the objective

$$\frac{1}{n}\sum_{j=1}^{n} w_j C_j$$

the greedy algorithm then yields a prioritized renumbering of the jobs $w_n/X_n \geqslant \ldots \geqslant w_1/X_1$ (this is Smith's or the $c\mu$-rule). Let Z^{G} and Z^{OPT} be the values of the objective associated with Smith's rule and with the optimal policy respectively. Then $Z^{\mathrm{OPT}} = O(n)$, whereas

$$\begin{aligned} Z^{\mathrm{G}} - Z^{\mathrm{OPT}} &\leqslant \frac{1}{n}\,\Phi(S_k)\frac{w_1}{X_1} + \sum_{k=2}^{n}\Phi(S_k)\left(\frac{w_k}{X_k} - \frac{w_{k-1}}{X_k}\right) \\ &= \frac{1}{n}\frac{w_n}{X_n}\frac{(M-1)^2}{M}\max_{j\in E}(X_j^2) = O\left(\frac{1}{n}\right). \end{aligned} \tag{66}$$

The $1 + O(1/n^2)$ performance ratio for batch indicates that Smith's rule is optimal except at the end of the schedule and the end effect is of the order $1/n$ per job.

Drawing an analogy with the case of scheduling an arrival stream of jobs with traffic intensity $0 < \rho < 1$, we might think that for each busy period the optimal holding costs per job would be of the order $1/(1-\rho)$, whereas the suboptimality per job would be of the order $1-\rho$. This would be in striking contrast with the order 1 suboptimality per job (or equivalently per unit time) which is obtained in Section 3. Further research is needed to discover whether these much stronger asymptotic optimality results indeed hold.

Colin McDiarmid (*University of Oxford*)
The mathematics of operational research contains various streams of endeavour, with different levels of interaction, which can create the unfortunate impression that it is not really a coherent subject. Two thriving and hitherto rather separate streams among these concern mathematical programming and the optimal control of stochastic systems. Methods and ideas from the former are imported in the 'achievable region' approach to the latter. It is pleasing and exciting to see these imports casting new light. The present contribution shows that this light is very powerful for certain classes of stochastic control problem and further indicates that the illumination may in the future extend more widely.

Rolf H. Möhring and Marc Uetz (*Technische Universität Berlin*) **and Andreas S. Schulz** (*Massachusetts Institute of Technology, Cambridge*)
The paper is an important recent enhancement of our ability to design and analyse policies for problems in queuing theory and stochastic control. The approach of using a relaxation of the performance space to derive controls with good performance is well established in deterministic optimization. It has evolved in recent years into a promising tool in stochastic optimization as well.

Most commonly used are linear programming (LP) relaxations, and, especially in the context addressed by the authors, generalizations of extended polymatroids. The latter give immediate rise to index policies as was nicely illustrated in Bertsimas and Niño-Mora (1996) for certain multiarmed bandit problems. The current climax of the approach discussed is its use to prove asymptotic optimality of a Gittins index policy for parallel servers (Section 3).

The 'achievable region approach' is closely related to the pre-existing paradigm of LP-based approximation algorithms. More specifically, work by for example Wolsey (1985), Queyranne (1993), Hall *et al.* (1997) and Möhring *et al.* (1998) has impressively revealed the potential of LP formulations in natural date variables to derive exact and approximate algorithms and characterizations of the space of feasible solutions in deterministic and stochastic scheduling.

In these settings the number of jobs is fixed, and we are interested in constant relative performance guarantees rather than asymptotic optimality. For the problem of minimizing the weighted total expected completion time $\Sigma\, w_j\, E(C_j)$ of n jobs on m identical machines with job processing times p_j fulfilling $V(p_j) \leqslant E(p_j)^2$, we have provided ρ-approximation algorithms with $\rho = 2 - 1/m$ and $\rho = 4 - 1/m$ for the cases without and with release dates respectively (Möhring *et al.*, 1998). More complicated approximation ratios ρ apply to general distributions.

During a visit by Glazebrook, we discussed our results and their relationship to the achievable region approach. He pointed out that our bounds yield asymptotic optimality in the case without release dates, and he also derived the aforementioned approximation ratio within the achievable region paradigm of approximate generalized conservation laws (GCLs). Subsequently, G. Weiss gave an alternative derivation of our inequalities from his analysis of Smith's rule (Weiss, 1992).

An important question concerns the handling of additional side-constraints such as precedence relationships among jobs. In this case, GCLs do not apply, but our LP-based approach still yields approximate results for the single-machine case; the case of m machines remains an interesting problem.

This paper demonstrates the power and applicability of the achievable region approach to queuing and other stochastic systems. Motivated by the success of LP-based approximations in combinatorial optimization in general, we wonder whether the LP-based paradigm that utilizes arbitrary LPs as relaxations of the achievable region will lead to progress in stochastic systems other than scheduling.

R. W. Owen (*University of Essex, Colchester*) **and M. M. Gregorio-Domínguez** (*Instituto Tecnologico Autonomo de Mexico, Mexico City*)

This paper is welcome not only for its novel uses of the achievable region approach but also for its clear exposition of the key concepts and features of the approach.

The results in expressions (15) and (35) may be combined to obtain a bound on the suboptimality of a Gittins index policy as a proportion of the optimum:

$$\frac{Z^{u_G} - Z^{\mathrm{OPT}}}{Z^{\mathrm{OPT}}} \leqslant \min\left(\frac{Z^{u_G} - Z^D}{Z^D}, \frac{\epsilon}{Z^{u_G} - \epsilon}\right)$$

where

$$\epsilon = \sum_{j=1}^{N} \Phi\{S(j)\}(G_j - G_{j-1}).$$

When an index policy is optimal, $\Phi \equiv 0$ and hence the bound will be tight.

We consider a single-machine scheduling problem in which a sequence-dependent switching time C_{ij} (with $E(C_{ij}) = c_{ij}$) is required to switch from job i to job j. The objective is to minimize the *flow time*, i.e. the sum of the completion times of all N jobs. An index policy is optimal for this problem when either all switching times are 0 or when the conditions in theorem 9 are satisfied. (Further details are given in Gregorio-Domínguez and Owen (1997).)

Theorem 9. If the jobs are numbered in increasing order of their expected processing times and the mean values of the corresponding switching times satisfy

(a) $c_{ij} \leqslant c_{ik}$ for $j < k$, $\forall i$,
(b) $c_{ki} \leqslant c_{ji}$ for $j < k < i$ and
(c) $c_{ij} \leqslant c_{ji}$ for $i < j$,

then the policy which orders the jobs in increasing order of their expected processing times minimizes in expectation the flow time in the class of non-pre-emptive policies.

Fig. 5 compares the exact percentage suboptimality with the bound on the percentage suboptimality given above. The various parts display the consequences of violating various combinations of conditions (a)–(c) in theorem 9.

As expected, the bound is tight when all conditions are satisfied and when all the processing times and switching times are identical (in which case any non-pre-emptive policy is optimal). Otherwise, the bound is generally far from tight.

The poor performance of the bound is put in perspective by observing that the flow time model with switching times is a time-dependent travelling salesman problem — an extreme problem on which to test any methodology.

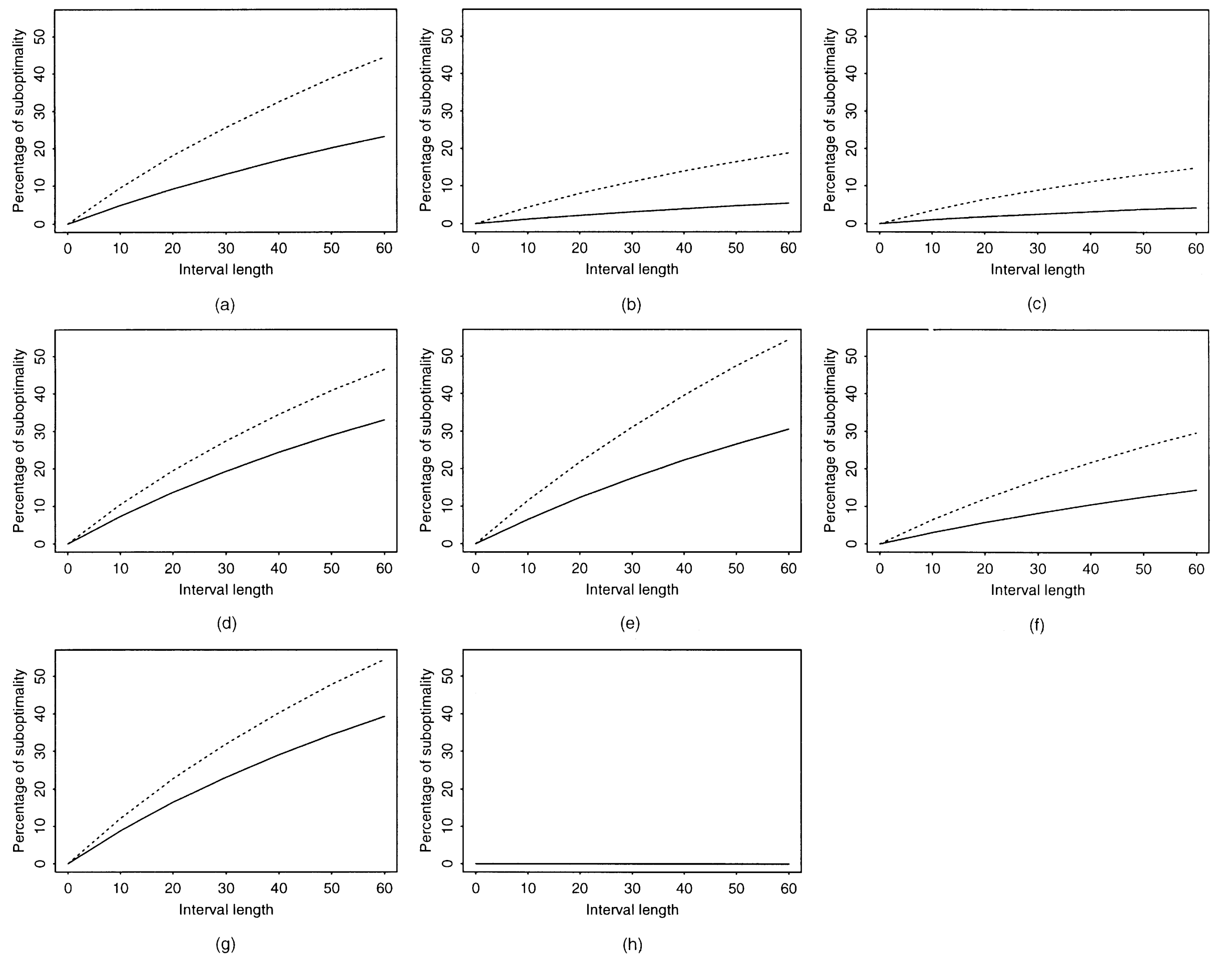

Fig. 5. Bound (·········) *versus* exact percentage suboptimality (———) ($N = 8$)—mean processing times are drawn independently from a uniform [50, b] distribution where b varies over the range [50, 110]—the switching time matrices are constructed to violate various combinations of conditions (a)–(c) of theorem 9: (a) condition (a) violated; (b) condition (b) violated; (c) condition (c) violated; (d) conditions (a) and (b) violated; (e) conditions (a) and (c) violated; (f) conditons (b) and (c) violated; (g) conditions (a)–(c) violated; (h) no conditions violated

J. Preater (*Keele University*)
The authors have given an illuminating account of the achievable region methodology. The approach is a most welcome relief from the strait-jacket of dynamic programming recursions: it will be interesting to see how many traditional dynamic programming areas will succumb to it.

One plausible forbear of the achievable region approach is multicriteria linear programming, where one considers, at least conceptually, the space of potential objective function vectors before reaching a consensus. It is apt therefore that a by-product of the new approach is that multicriteria versions of the problems are thereby made more transparent.

Ulrich Rieder (*University of Ulm*)
The control of stochastic and dynamic systems is a mathematically challenging problem. Conventional approaches are based on stochastic dynamic programming techniques. In recent years the 'achievable region approach' has been developed to formulate stochastic control problems (predominantly queuing systems) as mathematical programming problems. The vectors of performance variables are typically expectations. Often only the first moments of the relevant distributions play a role.

The authors have done some important work on the development of powerful new methodologies to analyse a range of stochastic optimization problems. In particular, the results for the difficult M-server queuing system are very interesting. The work is important because of its potential usefulness in many further applications, especially in stochastic dynamic optimization problems with side-constraints (as discussed at the end of Section 4). Indeed, knowing approximate generalized conservation laws (like those in Section 3) may give performance guarantees for simple bench-mark policies. Bounds for simple policies such as index policies and threshold policies are easily derived by the achievable region approach.

All the applications in the paper involve linear objectives expressed in terms of time averages. No statistics other than the first moments (taken with respect to the stationary distribution) define the linear programs. These assumptions and properties seem to be significant for the successful application of the achievable region approach. I wonder whether this approach could be extended to even more general classes of stochastic dynamic optimization problems, e.g. decentralized stochastic control problems or finite horizon portfolio problems.

Shaler Stidham, Jr (*University of North Carolina, Chapel Hill*)
The authors have set for themselves two goals:

(a) to provide a concise and widely accessible survey of the achievable region approach and
(b) to present new methods and results which they, among other researchers, have developed in recent years.

They have achieved both goals in this well-crafted paper, which should become a standard reference on the subject.

Although the ideas underlying the achievable region approach are simple, extracting the key concepts from the literature and presenting them in a clear and cogent manner, with simple but accurate notation, is not a trivial task. The authors are to be commended for making it look easy in Section 2.

The remaining sections of the paper cover recent extensions of the achievable region approach by the authors and colleagues. Section 5 contains results that I found particularly interesting. Previous research has applied the achievable region approach primarily to scheduling problems. The authors present preliminary results on using the achievable region approach to prove the optimality of a threshold policy for setting optimal arrival and service rates in a queuing facility. Previous research (including my own) on this topic has typically used inductive arguments based on a dynamic programming formulation of the problem. By contrast, the achievable region approach considers the vector of the two performance measures — throughput and inventory — in essentially the same way that the vector of performance measures in different classes is considered in scheduling problems. Although the achievable region approach apparently has not yet solved any new problems in this area, it is nonetheless a potentially useful addition to the analyst's toolkit.

The authors point out that, although most of the applications of the achievable region approach have been to problems in which the performance measures are steady state expectations, 'the approach is certainly not limited in its usefulness to such problems'. Recently we have shown that strong conservation laws hold for many problems on a finite time, sample path basis (Green and Stidham, 1999). Applications include scheduling in multiclass stochastic fluid systems and input–output systems in which the cumulative inputs and/or the cumulative (potential) output are mixtures of continuous and jump processes (e.g. Levy processes). Our approach exploits recent results concerning the solution to the Skorohod problem for such systems. We are currently extending these results to systems driven by a Brownian netput process and to multiclass networks.

Lyn Thomas (*University of Edinburgh*)
The authors are to be congratulated for showing so clearly how the mathematical programming approach to complex service systems can exploit the generalized conservation laws that the optimal operating strategies for such systems must satisfy. Although dynamic programming (DP) allows more flexibility in the form of the objective function, it is true that DP finds it difficult to exploit the fact that the optimal solution must be in such a limited class of policies.

However, I would like to comment on the potential for this mathematical programming approach to be used for game theory models of such queuing systems. The paper assumes that there is some overall controller of the queue who is interested in minimizing the total cost of running the system. An alternative problem is that the customers in the queue are all individual decision makers who make their own choices about how they move through the system and which servers they choose according to what is best for them. This is an n-person game and I am struck that the convex polyhedron P defined in equation (13) is exactly the definition for the core of a game with characteristic function $b(S)$. The corresponding game is not sensible without the introduction of benefits to the customers of passing through the system, but it does seem that the mathematical programming approach would fit in very well with the standard ways of solving such games. This may allow us to extend this work to deal with versions where each customer is an individually rational decision maker in much the same way as the work of Tijs (1992) and Feltkamp (1995) has extended deterministic scheduling problems. In their intriguing extensions of scheduling, owners of each job being scheduled are co-operating and competing to maximize their own individual benefits.

David D. Yao (*Columbia University, New York*)
The paper has done an excellent job in presenting an overview of the so-called ‘achievable region approach’ to a class of stochastic scheduling or control problems that are usually solved by dynamic programming techniques.

The starting-point of the approach is to characterize the performance space of a system under control by a set of linear constraints, a total of 2^n of them, n being the number of job classes. In general, even optimizing a linear objective function over such a feasible space is intractable (because of the exponential number of constraints).

The crucial point here is that certain terms on the right-hand side of the constraints could be ignored, resulting in a bound and, even better, a structure that is known as an extended polymatroid (EP). And, optimizing a linear objective over an EP is solvable by a greedy algorithm, which directly yields the Gittins indices, which, in turn, translate into a priority rule. Furthermore, relaxing the constraints creates a duality gap, which provides a natural bound for the performance, in terms of the gap between the objective value under the priority rule and the objective value under the optimal policy.

However, we still do not know what is the optimal solution, in terms of the control policy, to the original problem. In fact, we do not even know whether the problem is polynomially solvable— polynomial in n, that is. It would be worthwhile to carry out numerical investigations along these lines using interior point techniques, which are mostly primal–dual based. The class of problems with an EP or an EP-like structure often has a rich duality in the sense that the dual variables yield the Gittins indices. In this regard, they appear to be ideal candidates for interior point techniques.

One might argue that the optimal control to the original problem might not be of much interest, as it will not be as ‘clean’ as the priority rule that is at once optimal to the relaxed problem and ‘close’ to the optimal performance via an error bound. However, in applications, an interior point, which can be realized by policies such as Kleinrock’s delay cost schedules, may have far superior variance properties than the priority rule has.

The connection to directional convexity is another interesting point that may lead to other results and applications. For instance, instead of splitting incoming traffic as in the load balancing problem, we may be able to obtain results regarding the correlation between arrivals. Specifically, use directional convexity as a measure for the ‘strength’ of correlation, and investigate what is the effect on the queuing performance.

The **authors** replied later, in writing, as follows.

We are delighted that the paper has provoked so many thoughtful responses and are grateful to the discussants for their insights and stimulus to further reflection and work. It is not realistic to respond to all the points raised. We have tried to focus on some of the main themes.

Complex problems and relaxations of achievable region X

Dr Gibbens explicitly and other discussants by implication raise the issue of problems in which the achievable region X is difficult to identify. Certainly, for many complex models, it will be an unrealistic goal to characterize X fully by means of explicit linear or convex constraints. This is precisely the situation in the more mature mathematical programming approach to combinatorial optimization, which had led researchers to develop fruitful solution approaches based on strong yet tractable linear or convex programming relaxations. See the contribution by Professor Möhring, Professor Schulz and Mr Uetz. A few studies, including the parallel server model of Section 3, and the example discussed by Dr Gibbens, support the idea that relatively simple relaxations of X can yield provably good heuristic policies. Further, in at least two problem domains of which we are aware (namely multiclass queuing networks and restless bandits) we have prescriptions for developing sequences of ever more complex yet tight relaxations $P_1 \supseteq P_2 \supseteq \ldots \supseteq X$. See Bertsimas *et al.* (1995) and Bertsimas and Niño-Mora (1999). In these cases the order of an effective relaxation may yield insights into the complexity of the problem.

Before passing on to other related issues, we pause to note that it is not necessary to identify an exact X in situations where optimization over relaxation $P \supseteq X$ yields an optimal solution which can be shown to be in X. Theorem 7 exemplifies this in that an optimal policy is obtained (theorem 7, part (d)) before X is identified (theorem 7, part (e)). As indicated in Professor Stidham's comments, Green and Stidham (1998) utilize this fact in work which more fully exploits the sample path nature of the argument in Section 2 which yields expressions (4)–(6).

Professor Whittle's stimulating contribution raises the important question of what we look for in a relaxation. As implied above, we wish ideally to look to a relaxation to provide both a policy and a performance guarantee in the form of a suboptimality bound for it. Although the development of general approaches to the production of policies from relaxations is without doubt an important research priority, producing the heuristic policy for the model in Section 3 is not the difficult part. Although Professor Whittle's relaxation produces the index policy quite easily, so does an arguably simpler relaxed model in which a single server works at rate M. So far as we know, Professor Whittle's 'restless bandit' relaxation has yet to yield closed form performance guarantees for the M-server problem with M fixed. It is the production of effective performance guarantees for the index policy which for us is the major achievement of Section 3, rather than the heavy traffic optimality result which is a rather easy consequence. In this respect, we have some sympathy with Professor Bather's comments on asymptotic optimality. That said, we should underline the fact that the asymptotic optimality result to which Professor Whittle refers is substantially different from ours. For us, the assumption that M remains fixed is central.

Tightness of performance guarantees

The issue of the tightness of the performance guarantees for the Gittins index policy discussed in Section 3 is raised by several discussants (Dr Frostig, Professor Weiss, Dr Owen and Dr Gregorio-Domínguez). We state in the paper that we have not aimed at producing the best possible analysis, but the comments of the discussants now suggest that we should give more information about what is possible. Full details may be found in Glazebrook (1999). Rewrite equation (28) in the form

$$\sum_{i \in S} V_i^S x_i^u = b(S) + \Phi^u(S) \tag{67}$$

where $\Phi^u(S)$ is the sum of the first and third terms on the right-hand side. Introduce the set functions Φ_* and Φ^*, defined as

$$\Phi_*(S) = \inf_{u \in \mathcal{U}} \{\Phi^u(S)\}, \qquad S \subseteq E, \tag{68}$$

and

$$\Phi^*(S) = \sup_{u: S \to S^c} \{\Phi^u(S)\}, \qquad S \subseteq E, \tag{69}$$

where the supremum in equation (69) is over controls which give S priority over S^c. We can show that

$$Z^{u_G} - Z^{\text{OPT}} = \sup_{u \in \mathcal{U}} \left(\sum_{j=1}^{N} [\Phi^{u_G}\{S(j)\} - \Phi^{u}\{S(j)\}](G_j - G_{j-1}) \right) \tag{70}$$

$$\leqslant \sum_{j=1}^{N} [\Phi^{u_G}\{S(j)\} - \Phi_*\{S(j)\}](G_j - G_{j-1}) \tag{71}$$

$$\leqslant \sum_{j=1}^{N} [\Phi^*\{S(j)\} - \Phi_*\{S(j)\}](G_j - G_{j-1}). \tag{72}$$

The performance guarantees described in inequality (35) and corollary 1 (and we assume those alluded to by our discussants) are obtained as upper bounds on inequality (72). The bound in inequality (72) is usually more tractable analytically than expressions (70) or (71) but may be much less tight. In the example described by Dr Owen and Dr Gregorio-Domínguez, the class of policies is finite and the exact suboptimality can easily be recovered from equation (70). Given that they in effect appear to be assessing Smith's rule as a policy for a travelling salesman type of problem, we are pleasantly surprised by evidence provided by the discussants of the quality of the guarantee based on inequality (72). The issues raised in the final paragraph of the contribution by Dr Frostig and Professor Weiss may depend for their resolution on an ability to exploit expressions (70) and (71) effectively.

On a related matter, Dr Gittins raises the issue of known optimality results for parallel server problems. A few such results have emerged from the achievable region approach (see Shanthikumar and Yao (1992)). It would be an interesting challenge to determine whether Weber's (1982) results for flow time could emerge from an analysis which first established a work decomposition result as in equation (67) and then proceeded to demonstrate that the quantity on the right-hand side of equation (70) is 0.

Scope of the achievable region approach

Our rather tentative discussion of the scope of the achievable region approach in Section 6 has elicited a range of comments, as we hoped it would. Several discussants (including Dr Zervos, Dr Bäuerle and Professor Rieder) suggest that the approach may be limited in its usefulness outside the class of stochastic optimization problems involving time averages and/or linear objectives, or at least may be at its most natural there. We are not willing to be too pessimistic just yet. In addition to the work already published concerning successful applications of the achievable region approach to models with a discounted cost criterion (see, for example Bertsimas and Niño-Mora (1996, 1999) and Glazebrook and Garbe (1998)) and another study by Bhattacharya *et al.* (1995) whose prime focus is the maximization of a concave function of a performance measure related to a multiclass $M/GI/1$-queue with Bernoulli feed-back, two of the discussants (Dr Collins and Professor Stidham) helpfully point to the effective application of the achievable region approach to two very different problem areas, neither of which is concerned with steady state expectations. Interestingly, both of these concern problems with finite horizons—an area about which Professor Birge expresses some concerns. On more general finite horizon problems, we are encouraged by the work of Kumar and Meyn (1996) who showed that it may be possible to formulate a finite time linear programming relaxation which provides transient bounds on system performance from any initial condition.

Of course it would be foolish to pretend that there are not very substantial challenges ahead in extending the scope of the achievable region approach. Indeed, there are many very difficult problems to be solved which *do* involve time averages and linear objectives. One such is the scheduling control of multiclass queuing networks. It may be that the prime role of the approach here will ultimately be to provide bounds on achievable performance against which heuristic policies developed by other means may be assessed. Some modest progress to this end has already been made but a further development of our capacity to generate constraints on performance measures for complex processes is required.

So far as non-linear problems are concerned, Section 5 shows that the achievable region approach is well suited to deal with objectives which are non-linear in performance measures (expectations). In the context of the simple example discussed in Section 2, it is straightforward to replace the linear objective in equation (8) by $\phi(x_1, x_2)$ for some general concave ϕ. The optimizing performance will usually be a non-extreme point of X, realizable by a range of controls. These controls will plainly all share the same profile of expectations $\{E_u(N_1), E_u(N_2)\}$ but may have highly variable higher moment behaviour. In particular some policies will have much better behaviour than others from the perspective of $\{\text{var}_u(N_1), \text{var}_u(N_2)\}$. Such issues are raised by Dr Ansell and alluded to by Professor Yao. Motivated by such considerations, it seems to us that a more natural but considerably more formidable challenge concerns the optimization of the expectation of a non-linear function of the underlying process. See Dr Zervos's comments. The solution of the problem described by Dr Ansell can be regarded as a modest first step in this direction. Important research priorities highlighted by this work include

(a) an enhancement of the current approach based on polynomial potential functions for the development of constraints on performances which are higher order (and mixed) moments for queuing systems and
(b) the extension of semidefinite programming methodology to problems with non-linear objectives and/or constraints.

Miscellanea
We thought that it would assist readers if we clarified and/or commented on a few additional issues raised in the discussion.

(a) On Professor Bather's point, the assumption of pre-emptive controls in the example of Section 2 simplifies the argument at one or two points but is not critical. The effect of a restriction to non-pre-emptive policies would be to leave equation (4) unchanged, while increasing the right-hand sides of inequalities (5) and (6). The resulting achievable region is a line segment P' which is that part of P (see Fig. 1) which lies between two points A' and B' which represent the performances corresponding to non-pre-emptive implementations of $1 \rightarrow 2$ and $2 \rightarrow 1$ respectively. The resulting optimal control is a non-pre-emptive implementation of the $c\mu$-rule.
(b) Dr Collins asks for general principles for identifying and constructing generalized conservation laws. The main approaches used to date have been

    (i) the deployment of sample path work conservation arguments (as in the example of Section 2),
    (ii) the utilization of potential function techniques (as in Section 3) and
    (iii) arguments based on flow conservation in queuing systems.

    We understand least about the third of these approaches. In many ways, it seems more natural than approach (ii) but is challenging technically.
(c) We are grateful to Professor Birge and Professor Thomas for the suggestion that applications of the approach to distributed control and game theory models may be possible.
(d) In answer to Dr Gittins's query regarding algorithms for computing indices for his multipopulation random sampling models, the approach described in the paper would involve firstly constructing a finite state approximation to the relevant process (with given initial state) rather as in equation (6.13) and following text of Gittins (1989). An application of the adaptive greedy algorithm would then coincide with the 'largest to smallest' algorithm of Robinson (1982).
(e) We are grateful to Professor Yao for his suggestion regarding interior point methods. However, in the particular case of our parallel server model in Section 3 our computational experience suggests that the steady state distributions of the queue length process under an optimal control and under the index policy are usually so very close that they are unlikely to have substantially different variance properties.

References in the discussion

Ansell, P. S., Glazebrook, K. D., Mitrani, I. and Niño-Mora, J. (1999) A semidefinite programming approach for the optimal control of a multi-class single server queueing system with imposed second moment constraints. *J. Oper. Res. Soc.*, to be published.
Bertsimas, D. and Niño-Mora, J. (1996) Conservation laws, extended polymatroids and multi-armed bandit problems: a polyhedral approach to indexable systems. *Math. Oper. Res.*, **21**, 257–306.
——(1999) Restless bandits, linear programming relaxations and a primal-dual index heuristic. *Oper. Res.*, to be published.
Bertsimas, D., Paschalidis, I. C. and Tsitsiklis, J. N. (1995) Branching bandits and Klimov's problem: achievable region and side constraints. *IEEE Trans. Autom. Control*, **40**, 2063–2075.
Bhattacharya, P. P., Georgiadis, L. and Tsoucas, P. (1995) Problems of adaptive optimization in multiclass $M/GI/1$ queues with Bernoulli feedback. *Math. Oper. Res.*, **20**, 355–380.
Birge, J. R. and Louveaux, F. V. (1997) *Introduction to Stochastic Programming*. New York: Springer.
Collins, E. J. and McNamara, J. M. (1998) Finite-horizon dynamic optimisation when the terminal reward is a concave functional of the distribution of the final state. *Adv. Appl. Probab.*, **30**, 122–136.
Feltkamp, V. (1995) Cooperation in controlled network structures. *Thesis*. Catholic University of Brabant, Brabant.

Fleming, W. H. and Rishel, R. W. (1975) *Deterministic and Stochastic Optimal Control*. New York: Springer.
Gibbens, R. J. and Kelly, F. P. (1995) Network programming methods for loss networks. *IEEE J. Select. Areas Communs*, **13**, 1189–1198.
Gibbens, R. J. and Reichl, P. C. (1995) Performance bounds applied to loss networks. In *Complex Stochastic Systems and Engineering* (ed. D. M. Titterington). Oxford: Clarendon.
Gittins, J. C. (1979) Bandit processes and dynamic allocation indices (with discussion). *J. R. Statist. Soc.* B, **41**, 148–177.
——— (1989) *Multi-armed Bandit Allocation Indices*. New York: Wiley.
Glazebrook, K. D. (1999) On the scheduling of a general multiclass queueing system with parallel servers. *Technical Report*. Department of Statistics, University of Newcastle, Newcastle upon Tyne.
Glazebrook, K. D. and Garbe, R. (1998) On a new approach to the analysis of complex multi-armed bandits. *Math. Meth. Oper. Res.*, **48**, 419–442.
——— (1999) Almost optimal policies for stochastic systems which almost satisfy conservation laws. *Ann. Oper. Res.*, to be published.
Green, T. C. and Stidham, S. (1999) Sample-path conservation laws, with applications to scheduling queues and fluid systems. Submitted to *Questa*.
Gregorio-Domínguez, M. M. and Owen, R. W. (1997) On the performance of index policies for single-machine stochastic scheduling problems with switching penalties. *Working Paper 97-04*. Department of Mathematics, University of Essex, Colchester.
Hall, L. A., Schulz, A. S., Shmoys, D. B. and Wein, J. (1997) Scheduling to minimize average completion time: off-line and on-line approximation algorithms. *Math. Oper. Res.*, **22**, 513–544.
Harrison, J. M. and Pliska, S. R. (1981) Martingales and arbitrage in multi-period securities markets. *Stoch. Processes Appl.*, **20**, 381–408.
Hunt, P. J. and Laws, C. N. (1993) Asymptotically optimal loss network control. *Math. Oper. Res.*, **18**, 880–900.
Kumar, P. R. and Meyn, S. P. (1996) Duality and linear programs for stability and performance analysis of queueing networks and scheduling policies. *IEEE Trans. Autom. Control*, **41**, 4–17.
McNamara, J. M., Webb, J. N. and Collins, E. J. (1995) Dynamic optimisation in fluctuating environments. *Proc. R. Soc. Lond.* B, **261**, 279–284.
Möhring, R. H., Schulz, A. S. and Uetz, M. (1998) Approximation in stochastic scheduling: the power of LP-based priority rules. *Technical Report 595*. Fachbereich Mathematik, Technische Universität Berlin, Berlin.
Pliska, S. R. (1986) A stochastic calculus model of continuous trading: optimal portfolios. *Math. Oper. Res.*, **11**, 371–382.
Queyranne, M. (1993) Structure of a simple scheduling polyhedron. *Math. Progrmmng*, **58**, 263–285.
Robinson, D. R. (1982) Algorithms for evaluating the dynamic allocation index. *Oper. Res. Lett.*, **1**, 72–74.
Shanthikumar, J. G. and Yao, D. D. (1992) Multi-class queueing systems: polymatroidal structure and optimal scheduling control. *Oper. Res.*, **40**, 293–299.
Snell, J. L. (1952) Applications of martingales systems theorems. *Trans. Am. Math. Soc.*, **73**, 293–312.
Tijs, S. H. (1992) LP-games and combinatorial optimization games. *Cah. Cent. Etud. Rec. Oper.*, **34**, 167–186.
Weber, R. R. (1982) Scheduling jobs with stochastic processing requirements on parallel machines to minimise makespan or flowtime. *J. Appl. Probab.*, **19**, 167–182.
Weber, R. R. and Weiss, G. (1990) On an index policy for restless bandits. *J. Appl. Probab.*, **27**, 637–648.
Weiss, G. (1992) Turnpike optimality of Smith's rule in parallel machines stochastic scheduling. *Math. Oper. Res.*, **17**, 255–270.
Whittle, P. (1998) Restless bandits; activity allocation in a changing world. *J. Appl. Probab.* A, **25**, 287–298.
Wolsey, L. A. (1985) Mixed integer programming formulations for production planning and scheduling problems. *12th Int. Symp. Mathematical Programming, Massachusetts Institute of Technology, Cambridge, Aug.*