

# Management of demand-driven production systems

Mike Chen, *Student member, IEEE*, Richard Dubrawski, and Sean Meyn, *Fellow, IEEE*.

**Abstract**—Control-synthesis techniques are developed for demand-driven production systems. The resulting policies are time-optimal for a deterministic model, and approximately time-optimal for a stochastic model.

Moreover, they are easily adapted to take into account a range of issues that arise in a realistic, dynamic environment. In particular, control synthesis techniques are developed for models in which resources are temporarily unavailable. This may be due to failure, maintenance, or an unanticipated change in demand.

These conclusions are based upon the following development:

- (i) Workload models are investigated for demand-driven systems, and an associated *workload-relaxation* is introduced as an approach to model-reduction.
- (ii) The impact of hard constraints on performance, and on optimal control solutions is addressed via Lagrange multiplier techniques. These include deadlines and buffer constraints.
- (iii) Rules for choosing appropriate *safety-stocks* as well as *hedging-points* are described to ensure robustness of control solutions with respect to persistent disturbances, such as variability in demand and yield.

**Index Terms**—Scheduling, routing, queueing networks, inventory models, optimal control.

## I. INTRODUCTION

THIS paper concerns production management in a single production system. Even within this limited scope, there are numerous control decisions required for successful operation. These include, *work release*, *routing*, *sequencing*, *lot sizing*, *batching*, and *preventive maintenance scheduling*. These operating decisions are based on a range of performance metrics such as minimizing lead-time and inventory.

An application of current interest in both academia and industry is the semiconductor fabrication system. A modern plant may cost upward of several billion dollars. These systems are subject to high variations in demand, and frequent maintenance requirements result in significant internal plant variability. Figure 1 shows a simplified model to illustrate some of the aforementioned issues. In this example there are two external demands. Production management involves work release decisions, as well as scheduling and routing decisions for the various resources.

This paper is based upon work supported by the National Science Foundation under Award Nos. ECS 02 17836, ECS 99 72957 and DMI 00 85165. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation

S. M. and M. C. are with the Department of Electrical and Computer Engineering and the Coordinated Sciences Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A..

R. D. is a senior Software Engineer at Viasat in San Diego

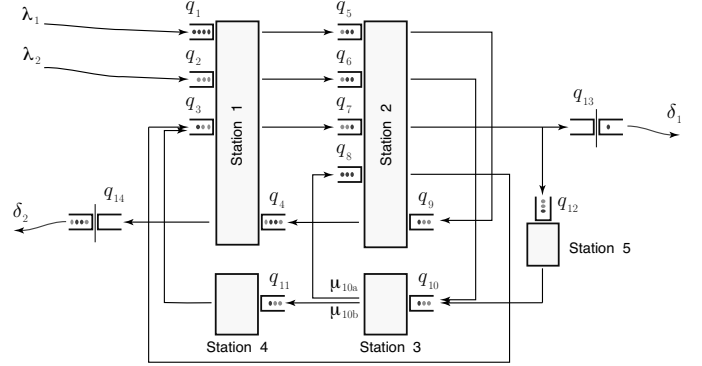


Fig. 1. A 16-buffer semiconductor production system.

Much of the literature on production management is based on optimal control, using either combinatorics or dynamic programming techniques. It is well known that either approach is typically computationally infeasible in realistic models. Moreover, even in cases where an optimal solution is computable, the solution may not provide qualitative insights on ‘closed-loop’ system behavior.

The policy synthesis techniques developed here are based upon simplified deterministic models, and relaxations of these based upon workload. We arrive at a geometric approach to policy synthesis for a deterministic fluid model, and translation to a model with variability is then performed using a combination of safety-stocks and generalized hedging points.

Many of these concepts extend or synthesize previous approaches to policy synthesis. In particular, the application of fluid models in policy synthesis has become widespread in recent years (see e.g. [4], [1], [24], [21]). Fluid models have been used recently to construct policies that minimize make-span, which is closely related to the time-optimality constraint imposed here [18], [10], [19].

The policies considered here are based on safety-stocks to avoid starvation of resources [5], and hedging-points to regulate inventory at cost-efficient locations [8], [12].

The workload-relaxation described in this paper is similar to the application of *state-space collapse* for networks in a heavy-traffic regime, where a stochastic model with Gaussian statistics is obtained via a Central Limit Theorem scaling [3], [14], [19]. The present paper does not concern limiting regimes, but instead applies the workload relaxation as a first step in policy synthesis.

A list of contributions of the present paper follows:

- (i) Workload-models are described for demand-driven production systems, and structural properties of the associ-

ated workload-relaxations are investigated.

- (ii) The *greedy-time-optimal* (GTO) policy is introduced in which time-optimality is utilized as a *constraint* in the deterministic model. It provides a path-wise optimal solution whenever a path-wise optimal solution exists. Moreover, the policy is computable using a finite-dimensional linear program.
- (iii) Sensitivity of performance and policy structure to *buffer constraints* is addressed using Lagrange-multiplier techniques.
- (iv) The GTO algorithm is adapted to account for scheduled maintenance, unanticipated breakdown, or transient exogenous demand.
- (v) The introduction of safety-stocks ensures that the policy is robust with respect to persistent, unpredictable disturbances.

Portions of these results have appeared in the thesis [11], and further numerical results may be found there.

The remainder of the paper is organized as follows. Section II contains background on fluid and workload models for demand-driven production systems, and develops various workload relaxations. Policy synthesis based on optimal control of these fluid models is considered in Section III. Section IV concerns policy synthesis techniques for systems subject to breakdown, maintenance, and fluctuations in demand and yield. Conclusions and plans for future research are contained in Section V.

N.B. IN 2003 A U.S. PATENT APPLICATION WAS APPROVED FOR A FAMILY OF POLICIES THAT INCLUDE THE COLLECTION OF GTO POLICIES DESCRIBED BELOW.

## II. NETWORK MODELS

We begin with a description of the network models to be considered, including stochastic and fluid network models, in addition to workload models and their relaxations. We also examine pull and push models, and describe how the former can be translated into the latter.

### A. Fluid network model

The *state process* in the fluid model, denoted  $\mathbf{q}$ , describes the flow of jobs at the buffers in the network. It evolves in a convex set  $\mathbf{X} \subseteq \mathbb{R}_+^{\ell_u}$ , and its trajectories are continuous and deterministic. The number of resources in the network is denoted  $\ell_m$ , and the number of *activities* (or controls variables) is denoted  $\ell_u$ .

For a given initial condition  $x^0 \in \mathbf{X}$ , the state process satisfies the time-evolution equations,

$$\mathbf{q}(t; x^0) = x^0 + B\mathbf{z}(t; x^0) + \alpha t, \quad t \geq 0. \quad (1)$$

The vector  $\alpha \in \mathbb{R}_+^{\ell_u}$  models exogenous arrival and departure rates, and  $B$  is an  $\ell_m \times \ell_u$  matrix that is defined by network topology, and long-run average routing and service rates.

The non-decreasing cumulative allocation process  $\mathbf{z}$  evolves on  $\mathbb{R}_+^{\ell_u}$ . For each  $1 \leq j \leq \ell_u$ ,  $\zeta_j(t; x^0) = \frac{d^+}{dt} z_j(t; x^0)$  is equal to the instantaneous percentage of time devoted to activity  $j$  at time  $t$ , where  $\frac{d^+}{dt}$  denotes the right-derivative.

The  $\ell_m \times \ell_u$  *constituency matrix*  $C$  has binary entries, with  $C_{ij} = 1$  if and only if activity  $j$  is performed using resource  $i$ . The allocation process is subject to the constraint  $\zeta(t; x^0) \in \mathbf{U}$ ,  $t \geq 0$ , where

$$\mathbf{U} := \{u \in \mathbb{R}_+^{\ell_u} : Cu \leq \mathbf{1}\}, \quad (2)$$

and  $\mathbf{1}$  denotes a vector consisting of ones. The condition  $Cu \leq \mathbf{1}$  indicates that resources are shared among activities, and that they are limited.

Throughout the paper we assume that the fluid model is *stabilizable*. That is, for each initial condition  $x^0 \in \mathbf{X}$ , one can find an allocation  $\mathbf{z}$  and a time  $T \geq 0$  such that  $\mathbf{q}(T; x^0) = x^0 + B\mathbf{z}(T) + \alpha T = \theta$ , where  $\theta \in \mathbb{R}^{\ell_u}$  is a vector of zeros. Characterizations of stabilizability are presented in [18], and some of these results are reviewed in Section III.

The fluid model may be motivated through a scaling of the following stochastic model:

$$Q(t; x^0) = x^0 - S(Z(t; x^0)) + R(Z(t; x^0)) + A(t), \quad (3)$$

where  $A, S$  are multi-dimensional renewal processes, and the allocation process  $Z$  is subject to the same constraints as  $\mathbf{z}$ .

While valuable in simulation and analysis, the stochastic model (3) has little value in policy synthesis when considered in this generality. Frequently, an effective policy may be generated through consideration of a simpler model that retains important details. We return to (3) in Section IV-D where we provide detailed statistical assumptions, and describe how to translate a policy from the fluid to the stochastic model.

In this paper we focus on *demand-driven* models (also known as *pull models*). An example is the 16-buffer network shown in Figure 1 in which demand rates for the two products produced are indicated by  $\{\delta_1, \delta_2\}$ . In a pull model it is assumed that demand is exogenous, while the arrival of new material is determined by the particular operating policy employed.

A virtual queue is used to model inventory for each product produced in a pull model: a negative value indicates deficit (unsatisfied demand), and a positive value indicates excess inventory. Ideally, a policy would regulate virtual queues to *zero*, though this is rarely feasible in practice. It is frequently desirable to maintain inventory through hedging-points to ensure that demand is satisfied with high probability.

We avoid the use of negative buffer levels, and instead translate a given pull model into an equivalent push model with state space  $\mathbf{X} \subset \mathbb{R}_+^{\ell_u}$ . We illustrate this construction using the network shown at left in Figure 2. This is a simple pull model with a single production resource, and a single virtual queue. At right is an equivalent push model, obtained by replacing the virtual queue by a single ‘exit-resource’ with two buffers. Raw material is modeled as the ‘supply resource’ with maximal rate  $\lambda$ . In the resulting push model, one buffer at the exit-resource is fed by the production resource, and the other is fed by the demand process with instantaneous rate  $\delta$ . These buffers are interpreted as surplus and deficit buffers, respectively. The

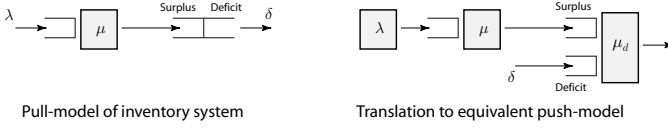


Fig. 2. Conversion of a pull model to a push model. If  $\mu_d \gg \mu$  then the two models shown are equivalent.

fluid model equation for the push model is given by (1) with

$$B = \begin{bmatrix} \lambda & -\mu & 0 \\ 0 & \mu & -\mu_d \\ 0 & 0 & -\mu_d \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0 \\ 0 \\ \delta \end{bmatrix}, \quad (4)$$

and the constituency matrix is the 3-by-3 identity matrix. Thus, after the push-to-pull conversion, a single-resource queue becomes a three-resource system.

Upon optimizing we will find that the policy at the exit-resource is non-idling, so that  $\zeta_3(t) = 1$  when  $q_2(t) > 0$  and  $q_3(t) > 0$ . In this case, the two systems are equivalent if  $\mu_d > 0$  is sufficiently large. Any push model may be transformed in this way to give a pull model evolving on  $\mathbb{R}_+^\ell$  for some integer  $\ell$ .

### B. Workload formulations and relaxations

Although considerably simpler than the stochastic network model, a fluid model may remain far too complex for exact analysis. In this section we introduce workload models and their relaxations. Frequently, these are of far lower complexity since the number of resources is typically far less than the number of buffers in a production system.

The definitions given here are taken from [19], some of which are based on related concepts in the heavy-traffic literature (e.g. [3], [14]).

Consider first the *velocity set*  $V$  given by

$$V := \{B\zeta + \alpha : \zeta \in U\}. \quad (5)$$

The fluid model may be described as a *differential inclusion* on  $X \subseteq \mathbb{R}_+^\ell$ , where the derivative of  $q$  is constrained to lie in  $V$ :

$$\frac{d^+}{dt} q(t; x^0) \in V, \quad q(t; x^0) \in X, \quad t \geq 0. \quad (6)$$

Since  $V$  is a polyhedron, it can be expressed as the intersection of half-spaces: for some integer  $\ell_v \geq 1$ , vectors  $\{\xi^i : 1 \leq i \leq \ell_v\} \subset \mathbb{R}^\ell$ , and constants  $\{\kappa_i : 1 \leq i \leq \ell_v\} \subset \mathbb{R}$ ,

$$V = \{v : \langle \xi^i, v \rangle \geq -\kappa_i, \quad 1 \leq i \leq \ell_v\}. \quad (7)$$

It follows that for a given  $x \in X$ , the minimal draining time is given by

$$T^*(x) = \max_{1 \leq i \leq \ell_v} \frac{\langle \xi^i, x \rangle}{\kappa_i}.$$

Stabilizability ensures that this is finite for  $x \in X$ .

Stabilizability also ensures that  $\theta \in V$ , so that  $\kappa_i \geq 0$  for each  $i$ . It is shown in [19] that these parameters have the specific form,

$$\kappa_i = \beta_i - \langle \xi^i, \alpha \rangle, \quad 1 \leq i \leq \ell_v,$$

where the  $\{\beta_i\}$  are binary-valued. We assume below that these are ordered so that, for some integer  $\ell_r \leq \ell_v$ ,  $\beta_i = 1$  for  $i \leq \ell_r$  and  $\beta_i = 0$  for  $i > \ell_r$ .

We thus arrive at the following definitions:

- (i) The vectors  $\xi^i$ ,  $1 \leq i \leq \ell_r$ , are called the *workload vectors*, and the index  $i$  is called the  $i$ th *pooled resource*.
- (ii) The vector load  $\rho \in \mathbb{R}^{\ell_r}$  is defined by,

$$\rho_i := \langle \xi^i, \alpha \rangle, \quad 1 \leq i \leq \ell_r,$$

and the system load is defined as the maximum,  $\rho := \max\{\rho_i\}$ .

- (iii) For a given state trajectory  $q$ , the corresponding *workload process* is defined by

$$w(t; w^0) = \Xi q(t; x^0), \quad (8)$$

where  $\Xi$  denotes the  $\ell_r \times \ell$  matrix with rows given by  $\{\xi^i : 1 \leq i \leq \ell_r\}$ , and  $w^0 = \Xi x^0$ .

The following assumptions are imposed throughout the paper.

- (A1) The network is stabilizable.
- (A2) The vector load entries  $\{\rho_i\}$  are non-increasing in  $i$ , and the system load satisfies  $\rho = \rho_1 < 1$ .
- (A3) There exists  $\ell_{ro} \leq \ell_r$  such that the workload vectors  $\{\xi^i : 1 \leq i \leq \ell_{ro}\}$  are *linearly independent*, and the minimal draining time may be expressed,

$$T^*(x) = \max_{1 \leq i \leq \ell_{ro}} \frac{\langle \xi^i, x \rangle}{1 - \rho_i}, \quad x \in X. \quad (9)$$

The relationship between system load and stabilizability is illustrated in Proposition 2.1. We say that the  $i$ th pooled resource is a *dynamic bottleneck* at time  $t$  if

$$T^*(q(t; x^0)) = \frac{\langle \xi^i, q(t; x^0) \rangle}{1 - \rho_i},$$

and we let  $\mathcal{I}(x)$  denote the index set of dynamic bottlenecks when  $q(t; x^0) = x$ . We obtain from these definitions the following proposition. For a proof, see [19, Proposition 2.5].

**Proposition 2.1:** Suppose that (A1)–(A3) hold. Then, for each initial condition  $x^0 \in X$ ,

- (a) For each  $x^1 \in X$ , provided this state is reachable from  $x^0$ , the minimum time to reach  $x^1$  from  $x^0$  is given by

$$T^*(x^0, x^1) := \max_{1 \leq i \leq \ell_v} \frac{\langle \xi^i, x^1 - x^0 \rangle}{\kappa_i}. \quad (10)$$

- (b) For any allocation-state pair  $(z, q)$ , the minimal draining time  $T^*(x^0)$  is achieved if and only if

$$\frac{d}{dt} \langle \xi^i, q(t; x^0) \rangle = -(1 - \rho_i),$$

for  $i \in \mathcal{I}(q(t; x^0))$  and a.e.  $0 \leq t < T^*(x^0)$ .

The dynamics of the workload process are *almost decoupled* under (A3):

**Proposition 2.2:** Under Assumptions (A1)–(A3) the workload process  $w$  is a differential inclusion with state space  $W := \{\Xi x : x \in X\}$ . For each  $t \geq 0$ , the following lower bounds hold:

$$\frac{d^+}{dt} w_i(t) \geq -(1 - \rho_i), \quad 1 \leq i \leq \ell_{ro}.$$

We now consider a model-relaxation in which the dynamics of the workload process are completely decoupled. For a given  $1 \leq n \leq \ell_{ro}$ , the  $n$ th workload relaxation is defined as follows:

(i) The velocity set is given by

$$\hat{V} = \{v : \langle \xi^i, v \rangle \geq -(1 - \rho_i), 1 \leq i \leq n\}. \quad (11)$$

Consequently, any solution  $\hat{q}$  to the  $n$ th relaxation evolves in  $X$ , and satisfies  $\frac{d^+}{dt}\hat{q}(t) \in \hat{V}$  for all  $t \geq 0$ .

(ii) The workload process for the relaxation is expressed

$$\hat{w}(t; w^0) = \hat{\Xi}\hat{q}(t; x^0), \quad t \geq 0, \quad (12)$$

where  $\hat{\Xi}$  is the  $n \times \ell$  matrix with rows equal to  $\{\xi^i : 1 \leq i \leq n\}$  and  $w^0 = \hat{\Xi}x^0$ . The workload process evolves on the workload space, given by

$$\hat{W} := \{\hat{\Xi}x : x \in X\}.$$

We have the following analog of Proposition 2.2. The proof is immediate from the definitions.

**Proposition 2.3:** Under Assumptions (A1)–(A3), for each  $1 \leq n \leq \ell_{ro}$ , the workload process  $\hat{w}$  for the the  $n$ th workload relaxation is a differential inclusion with state space  $\hat{W}$ , and velocity constraints given by

$$\frac{d^+}{dt}w(t; w^0) \in \{\phi \in \mathbb{R}^n : \phi_i \geq -(1 - \rho_i), 1 \leq i \leq n\}. \quad (13)$$

That is, the dynamics of the relaxed workload process  $\hat{w}(t; w^0)$  are decoupled.

Now that we have specified the models to be considered, we turn to the control synthesis problem in this deterministic setting.

### III. POLICY SYNTHESIS

The policies considered in this paper are based on the consideration of a cost function  $c: \mathbb{R}_+^\ell \rightarrow \mathbb{R}_+$ . We assume that  $c$  is piecewise linear, of the specific form,

$$c(x) = \max(\langle c^i, x \rangle : i = 1, \dots, \ell_c) \quad (14)$$

with  $c^i \in \mathbb{R}^\ell, i = 1, \dots, \ell_c$ . We further assume that  $c$  is non-negative on  $\mathbb{R}_+^\ell$ , and vanishes only at the origin.

For a given cost function we consider optimal control formulations for the fluid workload model and its relaxations. This requires a translation of cost in buffer coordinates to the *effective cost* on the space of workload configurations.

#### A. The effective cost

For the  $n$ th relaxation, we say that two states  $x, y \in X$  are *exchangeable* if  $\hat{\Xi}x = \hat{\Xi}y$ . Equivalently,  $\hat{T}^*(x, y) = \hat{T}^*(y, x) = 0$ , where

$$\hat{T}^*(x^0, x^1) := \max_{1 \leq i \leq n} \frac{\langle \xi^i, x^1 - x^0 \rangle}{1 - \rho_i}, \quad x^0, x^1 \in X.$$

For the purposes of optimization, if  $\hat{q}(t; x^0) = y$  and there is an exchangeable state  $x$  satisfying  $c(y) > c(x)$ , then there is no reason to remain at the state  $y$ . This motivates the following definitions for the  $n$ th relaxation:

**EFFECTIVE COST:**  $\bar{c}: \hat{W} \rightarrow \mathbb{R}_+$  is defined for  $w \in \hat{W}$  as the value of the linear program,

$$\begin{aligned} \bar{c}(w) := \min \quad & \eta \\ \text{s. t.} \quad & \langle c^i, x \rangle \leq \eta, \quad 1 \leq i \leq \ell_c, \\ & \hat{\Xi}x = w, \quad x \in X, \quad \eta \in \mathbb{R}_+. \end{aligned} \quad (15)$$

Since  $c$  is piece-wise linear it follows that this is also true for the effective cost,

$$\bar{c}(w) = \max\{\langle c^i, w \rangle : 1 \leq i \leq \ell_c\}, \quad (16)$$

where  $\{c^i : i \leq \ell_c\} \in \mathbb{R}^n$  are the extreme points obtained in the dual of (15).

**EFFECTIVE STATE:** This is any optimizer in (15),

$$\mathcal{X}^*(w) := \arg \min_{x \in X} (c(x) : \hat{\Xi}x = w). \quad (17)$$

**OPTIMAL EXCHANGEABLE STATE:** For each  $x \in X$ ,

$$\mathcal{P}^*(x) := \mathcal{X}^*(\hat{\Xi}x). \quad (18)$$

**MONOTONE REGION:** The subset of  $\hat{W}$  on which the effective cost is monotone,

$$\hat{W}^+ = \left\{ w \in \hat{W} \text{ such that } \bar{c}(w) \leq \bar{c}(w') \text{ whenever } w' \geq w, \text{ and } w' \in \hat{W} \right\}. \quad (19)$$

The effective cost is called *monotone* if  $\hat{W}^+ = \hat{W}$ .

In practice, the state space is subject to upper bounds on buffer capacity. It is important to understand the impact of such constraints on performance and policies. Quantifying *sensitivity* of performance with respect to buffer constraints is of particular interest in network planning.

Let  $b \in \mathbb{R}_+^\ell$  denote the  $\ell$ -dimensional vector of buffer constraints, satisfying  $0 < b_i \leq \infty$  for  $1 \leq i \leq \ell$ , and set  $X := \{x \in \mathbb{R}_+^\ell : x \leq b\}$ . The effective cost subject to these buffer constraints is obtained in the following linear program (the dual of (15)):

$$\begin{aligned} \max \quad & \gamma^T w - \beta^T b \\ \text{s. t.} \quad & \hat{\Xi}^T \gamma - \beta \leq c, \quad \beta \geq \theta. \end{aligned} \quad (20)$$

The variable  $\gamma$  in (20) is not sign-constrained.

The optimizers  $(\gamma^*, \beta^*)$  to (20) are Lagrange multipliers. Consequently,  $\gamma^*$  provides sensitivity of the effective cost to workload, and  $\beta^*$  provides sensitivity with respect to buffer constraints. The following result is a consequence of this observation, and the fact that an optimal solution to (20) may be found among *basic feasible solutions*. We denote these by  $\{(\gamma^i, \beta^i) : 1 \leq i \leq \ell_{\bar{c}}\}$  (note that the integer  $\ell_{\bar{c}}$  defined here is in general larger than the integer used in (16)).

**Proposition 3.1:** For the  $n$ th relaxation, and each  $w \in \hat{W}$ ,

(a) The effective cost  $\bar{c}(w) = \bar{c}(w; b)$  is given by

$$\bar{c}(w) = \max_{1 \leq i \leq \ell_{\bar{c}}} (w^T \gamma^i - b^T \beta^i). \quad (21)$$

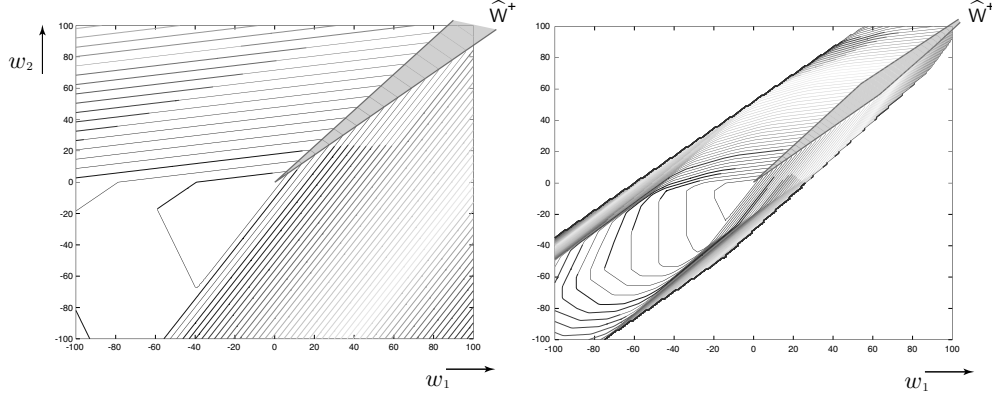


Fig. 3. Two instances of the effective cost  $\bar{c}$  for the two-dimensional relaxation of the network shown in Figure 1. On the left is shown contour plots of the effective cost when no buffer-constraints are imposed; at right is shown contour plots of the effective cost with buffer-constraints given in (24). In both figures, the shaded regions are monotone.

(b) Suppose that there is a unique maximizing index in (21), denoted  $i^* = i^*(w, b)$ . Then,

$$\begin{aligned} \frac{\partial}{\partial w_j} \bar{c}(w; b) &= \gamma_j^{i^*}, & 1 \leq j \leq n; \\ \frac{\partial}{\partial b_j} \bar{c}(w; b) &= -\beta_j^{i^*}, & 1 \leq j \leq \ell. \end{aligned}$$

We illustrate Proposition 3.1 using a two-dimensional relaxation of the 16-buffer model shown in Figure 1. The demand rate for each of the two products is equal to 19/75, and the vector of service rates is given by

$$\mu = (13/15, 26/15, 13/15, 26/15, 1, 2, 1, 2, 1, 1/3, 1/2, 1/10, 100, 100)^T. \quad (22)$$

The cost on the buffers is linear, with

$$c = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 5, 5, 10, 10)^T. \quad (23)$$

A contour plot of the effective cost without buffer constraints is shown at left in Figure 3. The figure shows that the monotone region  $\hat{W}^+$  is a small subset of the entire feasible region  $\hat{W}$ .

The level sets change dramatically when buffer constraints are imposed. Consider the vector of constraints given by

$$b = (10, 10, 20, 20, 20, 10, 10, 10, 10, 10, 10, 30, 30, 30, 30)^T. \quad (24)$$

The effective cost is now computed via (21). The level sets of  $\bar{c}$  are shown on the right hand side of Figure 3.

Based on this geometry we find that it is not difficult to devise efficient policies for a workload model of moderate dimension. Several general methods are developed in the following sections. The question then arises, *how can this be adapted to provide a policy for the original complex network of interest?*

This question is addressed in Section 4 of [19] for models in heavy traffic, but the same methods may be adapted to the present setting.

Suppose that  $[\hat{q}^\circ(t; x^0), \hat{z}^\circ(t; x^0)]$  is a piecewise linear, feasible solution to the  $n$ th relaxation. We assume that  $\hat{q}^\circ(t; x^0) = \mathcal{P}^*(\hat{q}^\circ(t; x^0))$  for all  $t > 0$ . Based on this solution for the relaxation, an allocation process  $\zeta$  is defined for the unrelaxed model as follows where, for each  $t \geq 0$ , the rate  $\zeta(t; x^0)$  is a function of  $[\hat{q}(t; x^0), \hat{\zeta}(t; x^0), q(t; x^0)]$ .

Given the current states  $y = q(t; x^0)$ ,  $y^\circ = \hat{q}^\circ(t; x^0)$ , define  $\mathcal{I}_c(y) := \{1 \leq j \leq \ell_c : c(x) = \langle c^j, y \rangle\}$ , and define  $\zeta(t; x^0) \in \mathcal{U}$  as any optimizer  $\zeta^* \in \mathcal{U}$  in the following linear program:

$$\begin{aligned} \min \quad & \gamma \\ \text{s. t.} \quad & \gamma \geq \langle c^i, v \rangle, \quad i \in \mathcal{I}_c(y) \\ & B\zeta + \alpha = v, \\ & v_i \geq 0, \quad \text{if } y_i = 0, \\ & v_i \leq 0, \quad \text{if } y_i = b_i, \quad 1 \leq i \leq \ell \\ & \langle \xi^j, v \rangle \leq \langle \xi^j, (B\hat{\zeta}^\circ + \alpha) \rangle, \\ & \quad \text{if } \langle \xi^j, y \rangle = \langle \xi^j, y^\circ \rangle, \quad 1 \leq j \leq n, \\ & \gamma \in \mathbb{R}_+, \zeta \in \mathcal{U}, v \in \mathbb{R}^\ell. \end{aligned}$$

These constraints ensure that  $w_i(t; x) \leq \hat{w}_i^\circ(t; x)$  for all  $i \leq n$  and all  $t$ . Consequently, this policy minimizes the instantaneous decrease in  $c(q(t))$  at each time instant, subject to the constraint that  $w$  never exceed  $\hat{w}$ .

The policy is stabilizing under general conditions on the model and policy. In [19, Theorem 4.1] asymptotic optimality is established for networks in heavy traffic when the effective cost is monotone.

We now show how to generate policies for the relaxation through optimal control techniques.

### B. Optimal control

In the fluid model in which there is no variability we consider the transient control problem: given an initial condition  $x^0 \in X$ , we wish to find an allocation  $z$  that drains the network in an economical manner. Below are two formulations of optimal control:

**TIME-OPTIMAL CONTROL:** For each initial condition  $x^0 \in X$ , find an allocation  $z$  that achieves the *minimal emptying time*, so that the resulting state trajectory  $q$  satisfies,

$$q(t; x^0) = \theta, \quad t \geq T^*(x^0).$$

**INFINITE-HORIZON OPTIMAL CONTROL:** For each initial

condition  $x^0 \in X$ , find an allocation  $z$  that minimizes the total cost,

$$J(x^0) = \int_0^\infty c(q(t; x^0)) dt. \quad (25)$$

We let  $J^*(x^0)$  denote the ‘optimal cost’, i.e., the infimum over all policies when the initial condition is  $x^0$ .

Under Assumption (A1), time-optimal and infinite-horizon optimal control solutions exist for each initial condition, although the solutions may not be unique.

These control-criteria extend to any workload relaxation. For the  $n$ th relaxation we may restrict attention to the workload process  $\hat{w}$  by exploiting the exchangeability of states  $x, y \in X$  with a common workload value. The optimization criteria are then defined with respect to the workload space  $\hat{W} \subset \mathbb{R}^n$ , and the cost function on  $\hat{W}$  is taken to be the effective cost.

If the effective cost is monotone, then the infinite-horizon optimal policy is *work conserving*, in the sense that  $\frac{d}{dt} \hat{w}(t; w^0) = -(1 - \rho)$  when  $\hat{w}(t; w^0)$  lies in the interior of  $\hat{W}$ . When monotonicity fails, so that  $\hat{W}^+$  is a proper subset of  $\hat{W}$ , then from certain initial conditions one may have  $\frac{d}{dt} \hat{w}_i(0; w^0) = +\infty$  for some  $1 \leq i \leq n$  in an optimal solution [19]. Consequently, the infinite-horizon optimal solution  $\hat{w}^*$  evolves for  $t > 0$  in a region  $\mathcal{R}^*$  satisfying

$$\hat{W}^+ \subset \mathcal{R}^* \subset \hat{W}$$

The boundaries of  $\mathcal{R}^*$  are interpreted as switching curves defining the optimal policy (see [15], [21], [6]).

Consider for example a two-dimensional relaxation of the 16-buffer model. We have  $\mathcal{R}^* = \hat{W}^+$  when  $(1 - \rho) \in \hat{W}^+$ . When this is not the case, a trade-off must be made between reducing the cost at time  $t = 0+$ , and draining the network in a timely manner. Consequently,  $\mathcal{R}^*$  is strictly larger than the monotone region  $\hat{W}^+$ . This is illustrated at left in Figure 4.

### C. Policies based on time-optimality

We are now prepared to introduce a class of policies that form the key building block for all of the policies considered below.

Although the infinite-horizon optimal policy for the fluid model is well motivated by the solidarity between optimal control solutions for fluid and stochastic models [17], [18], [19], [6], [20], in practice one may consider related policies with desirable properties that are more easily computed.

As motivation, consider again the two-dimensional relaxation of the 16-buffer model. Suppose that the vector  $(1 - \rho)$  lies below  $\hat{W}^+$ , as illustrated at left in Figure 4. Observe that the lower boundary of  $\mathcal{R}^*$  is defined by a linear switching curve  $s_1^*(w_1)$  satisfying

$$bw_1 < s_1^*(w_1) < aw_1, \quad w_1 > 0, \quad (26)$$

where  $b = (1 - \rho_2)/(1 - \rho_1)$ , and  $a$  is the slope of the lower boundary of the monotone region  $\hat{W}^+$ . Consequently, for initial conditions lying below the infinite-horizon optimal switching curve with slope  $s_1^*$ , the infinite-horizon optimal control is not time-optimal.

The *greedy* (or *myopic*) policy defines the allocation rate  $\hat{\zeta}(t)$  at time  $t$  so that  $\frac{d}{dt} c(\hat{q}(t; x^0)) = \frac{d}{dt} \bar{c}(\hat{w}(t; w^0))$  is minimized. Consequently, in the workload relaxation, for any initial  $w^0 \in \hat{W}$ , the workload trajectory satisfies  $\hat{w}(t; w^0) \in \hat{W}^+$  for all  $t > 0$ .

One obtains a time-optimal allocation if the region  $\mathcal{R}^*$  is expanded to the set  $\mathcal{R}^{\text{GTO}}$  shown on the right hand side of Figure 4 since we then have  $(1 - \rho) \in \mathcal{R}^{\text{GTO}}$ . This is one example of the *greedy-time-optimal* (GTO) policy described next.

**GTO POLICY** This is the state-feedback policy, defined for the  $n$ th relaxation as follows. Given the current state  $\hat{q}(t; x^0) = x$ , the allocation rate  $\hat{\zeta}(t) \in U$  is defined to be any optimizer in the linear program

$$\begin{aligned} \min \quad & \eta \\ \text{s. t.} \quad & \langle c^i, v \rangle \leq \eta, \\ & \quad \text{if } i \in \mathcal{I}_c(x), 1 \leq i \leq \ell_c, \\ & \langle \xi^i, v \rangle = -(1 - \rho_i), \\ & \quad \text{for } i \in \mathcal{I}(x), 1 \leq i \leq n, \\ B\zeta + \alpha &= v, \\ v_i &\geq 0, \text{ if } x_i = 0, 1 \leq i \leq \ell, \\ v_i &\leq 0, \text{ if } x_i = b_i, 1 \leq i \leq \ell, \\ & \zeta \in U, v \in \mathbb{R}^\ell, \eta \in \mathbb{R}. \end{aligned} \quad (27)$$

In words, the GTO policy minimizes the derivative  $\frac{d}{dt} c(\hat{q}(t))$ , subject to time-optimality, and state space constraints.

The following properties of the GTO policy follow directly from Proposition 2.1 and the definitions. These conclusions hold for the queueing model with velocity space  $V$ , as well as any of its workload-relaxations.

**Proposition 3.2:** Let  $\hat{q}(t; x^0)$  denote a trajectory defined by a GTO policy with initial condition  $x^0$ . Then,

(a) The state trajectory  $\hat{q}$  is time optimal. That is,

$$\hat{q}(t; x^0) = \theta, \quad t \geq T^*(x^0).$$

(b) Suppose that the linear program (27) has a unique solution for each  $x \in X$ . If a unique path-wise optimal solution  $\hat{q}^*$  exists starting from  $x^0$  then  $\hat{q}(t; x^0) = \hat{q}^*(t; x^0)$  for  $t \geq 0$ .<sup>1</sup>

(c)  $c(\hat{q}(t; x^0))$  is convex and non-increasing in  $t$ .

(d)  $\mathcal{I}(\hat{q}(t; x^0))$  is non-decreasing in  $t$ .

## IV. DECISION MAKING IN A DYNAMIC ENVIRONMENT

We now develop extensions of the GTO policy to address a range of issues that arise in a dynamic environment. In particular, we find that the GTO policy may be adapted to provide effective policies in the following circumstances:

- (i) A *transient* demand is imposed, and is given priority over other products in the system.
- (ii) Some component in the production network is temporarily in-operable, resulting in down-time for resources in the network. During repair it is still necessary to choose

<sup>1</sup>A solution is path-wise optimal if for *each* time instant,  $\hat{q}^*$  minimizes  $c(q(t; x^0))$  over *all* feasible trajectories  $q$ .

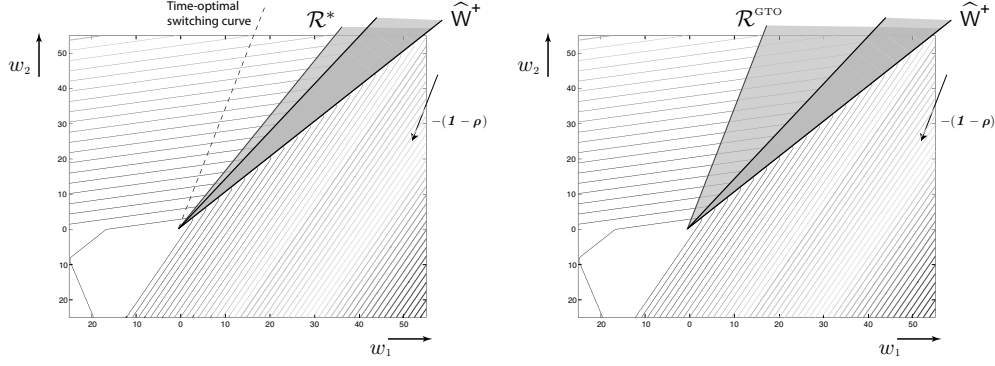


Fig. 4. On the left is shown the switching region for the infinite-horizon optimal policy,  $\mathcal{R}^*$ , and on the right is shown the switching region  $\mathcal{R}^{\text{GTO}}$  for the GTO policy. In this numerical example, the vector  $(1 - \rho)$  does not lie in  $\widehat{W}^+$ . Consequently, the region  $\mathcal{R}^*$  is strictly larger than the monotone region  $\widehat{W}^+$ , and  $\mathcal{R}^{\text{GTO}}$  is strictly larger than  $\mathcal{R}^*$ .

allocations at those resources in the network that are functioning.

- (iii) Some resources require preventative maintenance, so that some portion of the network is disabled for a period of time in the future. In this case, decisions regarding allocations prior to maintenance will be made subject to the knowledge of approaching down-time, and subsequent maintenance.
- (iv) The network is subject to persistent, unpredictable disturbances, such as uncertainty in demand and yield.

The impact of these disturbances can be reduced if the control synthesis problem is solved using all relevant information.

Until Section IV-D, we ignore persistent disturbances due to uncertainty in demand and yield. For previous investigations on persistent machine failures, see [23].

#### A. Hot-lots

In production parlance, a *lot* is a group of products in the system. A *hot-lot* is a lot (or group of lots) that is given high priority. It may be a prototype for a new product, in which case the (long-run) demand rate will be zero. We assume throughout that this is the case. Therefore, for the purpose of policy synthesis, a hot-lot is modeled in the initial condition of the network through the introduction of a corresponding virtual queue.

Priorities may be hard or soft. We consider three general classes below:

**NORMAL-PRIORITY HOT-LOT** The deficit buffer for the hot-lot is non-empty at time  $t = 0$ , and the holding cost at this deficit buffer is commensurate with the holding cost at other deficit buffers in the network.

**HIGH-PRIORITY HOT-LOT** The deficit buffer for the hot-lot is non-empty at time  $t = 0$ , and the holding cost at this deficit buffer is far larger than the holding cost at other deficit buffers.

**CRITICAL-PRIORITY HOT-LOT** The deficit buffer for the hot-lot is non-empty at time  $t = 0$ , and the allocation is chosen to meet this demand in the shortest time possible.

The high-priority hot-lot is used to model a finite size customer order where, due to the premium paid by the customer,

the system is realigned to meet the order quickly; a critical-priority hot-lot receives the highest priority possible.

The GTO policy may be applied without modification in systems with normal-priority or high-priority hot-lots since positive demand-rates are not required in the definition (27).

Policy synthesis for a critical hot-lot is more complex. One must first obtain the minimal clearing time of the hot-lot. Based on this, the minimal draining time for the overall system is computed, given that the critical hot-lot is cleared in minimal time. Each of these computations may be cast as a finite-dimensional linear program. The required modification in (27) is obvious and will not be presented here.

To illustrate the impact of a hot-lot on overall system performance we consider a version of the network model shown in Figure 1 in which the demand rate  $\delta_2$  is set to zero. However, the corresponding buffer  $q_{16}$  is non-zero since we assume that there is a transient demand for this product. In the simulation below the initial buffer levels are given by,

$$x^0 = (0, 5, 8, 0, 0, 4, 6, 6, 0, 7, 2, 2, 0, 0, 15, 20)^T. \quad (28)$$

Note that all of the buffers in the hot-lot path are initially empty. Consequently, to meet the hot-lot demand  $q_{16} = 20$  it is necessary to bring into the entrance buffers all of the required raw material.

We first consider a simulation to establish a baseline for analysis. The hot-lot is absent, so that  $q_{16}(0) = 0$ , and the remaining initial-condition values are unchanged in (28). Results from the GTO policy are shown at left in Figure 5. This illustrates normal operation of the network under the GTO policy when there is a single recurrent demand with rate  $\delta_1$ .

Consider now a normal-priority hot-lot with initial condition (28), and linear cost function defined in (23). A simulation of results obtained using the GTO policy is shown at right in Figure 5. The emptying time  $T^*(x^0) \approx 160$  in this simulation is approximately equal to the emptying time for the baseline system. The hot-lot is not cleared until late into the time-horizon  $[0, T^*(x^0)]$ .

The plot on the left hand side of Figure 6 shows a simulation of the network with a high-priority hot-lot under the GTO policy in which the holding cost at buffer 16 is increased from 10 to  $10^4$ , with initial condition fixed given in (28).

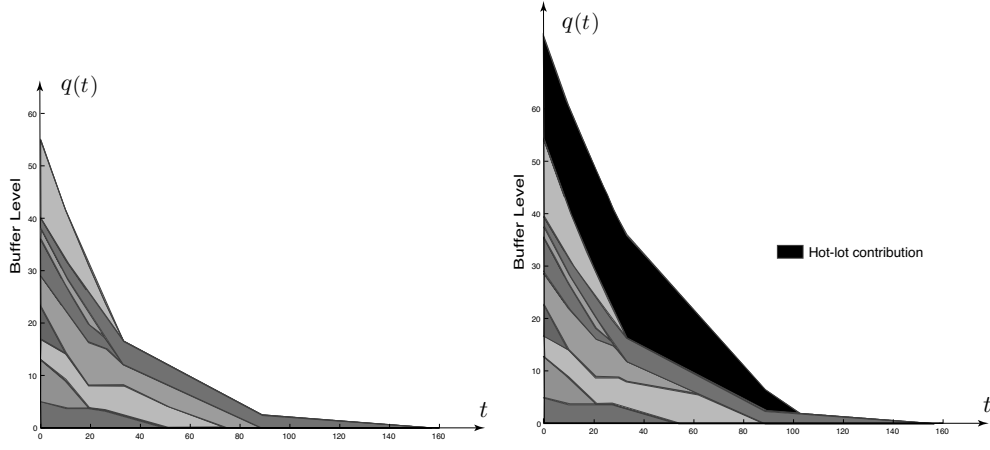


Fig. 5. The plot on the left hand side shows buffer levels versus time for normal operation under the GTO policy. The plot on the right hand side shows buffer levels versus time for a normal-priority hot-lot of size 20.

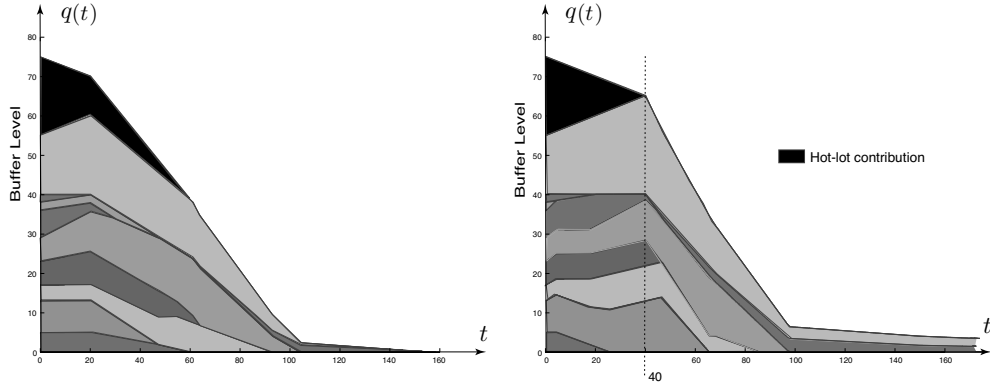


Fig. 6. On the left is a plot of buffer levels for the network shown in Figure 1 with a high-priority hot-lot under the GTO policy. On the right is a plot of buffer levels versus time under the GTO policy for the same network when a critical hot-lot is present. The system clearing-time was found to be  $T^* = 325$  in this case, which is approximately twice the minimal clearing time observed in the previous three simulations.

The resulting state trajectory clears the hot-lot demand in approximately 60 time units as opposed to about 100 time units for the normal-priority hot-lot. Of course, because the GTO policy imposes a global time-optimality constraint, the system drains at time  $T^*(x^0) \approx 160$ , exactly as seen for the normal-priority hot-lot.

Finally, we consider a critical-priority hot-lot. The minimal time required to clear the hot-lot deficit buffer is 40 time units with this initial condition. The minimal draining time for the network, subject to the constraint that the critical-priority deficit is cleared at time  $t = 40$ , is equal to approximately 325. This is approximately twice the clearing time seen for the normal or high-priority model. A simulation is shown at right in Figure 6.

### B. Unanticipated breakdown

Unscheduled down-time was ranked as the most significant cause of capacity loss in semiconductor fabs according to [22]. If a critical resource is lost without warning, large deficits can be generated while both upstream and downstream resources are forced into idleness.

We introduce here an extension of the GTO policy for network regulation during a period in which some resource is unavailable. Our goal is to obtain allocations that minimize

the impact of these gross disturbances. We find in simulations that the proposed policy places the network in a position to return to its normal operating state quickly once the resource is again available.

By normalization we may assume that the breakdown occurs at time  $t = 0$ , and take  $x^0$  as the state at this time. It is assumed that the time to repair, denoted  $T_{\text{MTTR}}$ , is known exactly. The acronym MTTR stands for *mean time to repair*, reflecting the fact that in practice only mean-values, and perhaps some higher order statistics, will be available.

Given that  $q(t; x^0) = y^0$  at some time  $0 \leq t < T_{\text{MTTR}}$ , the minimum time required to empty the system  $T^*(y^0; T_{\text{MTTR}}, t)$  is the value of the following linear program:

$$\begin{aligned} \min \quad & T \\ \text{s. t.} \quad & y^1 = y^0 + (B^{\text{down}} \zeta^1 + \alpha) \cdot T_1, \\ & \theta = y^1 + (B \zeta^2 + \alpha) \cdot T_2, \\ & \zeta^1 \in \mathcal{U}^{\text{down}}, \zeta^2 \in \mathcal{U}, y^1 \in \mathcal{X}, \end{aligned} \quad (29)$$

where  $T_1 := T_{\text{MTTR}} - t$ ;  $T_2 := T - T_{\text{MTTR}}$ ;  $\mathcal{U}^{\text{down}}$  is the set of feasible control rates during the period  $[0, T_{\text{MTTR}}]$  when a resource is disabled; and  $B^{\text{down}}$  is the corresponding system matrix in (1). When  $t > T_{\text{MTTR}}$  we may define  $T^*$  using (9).

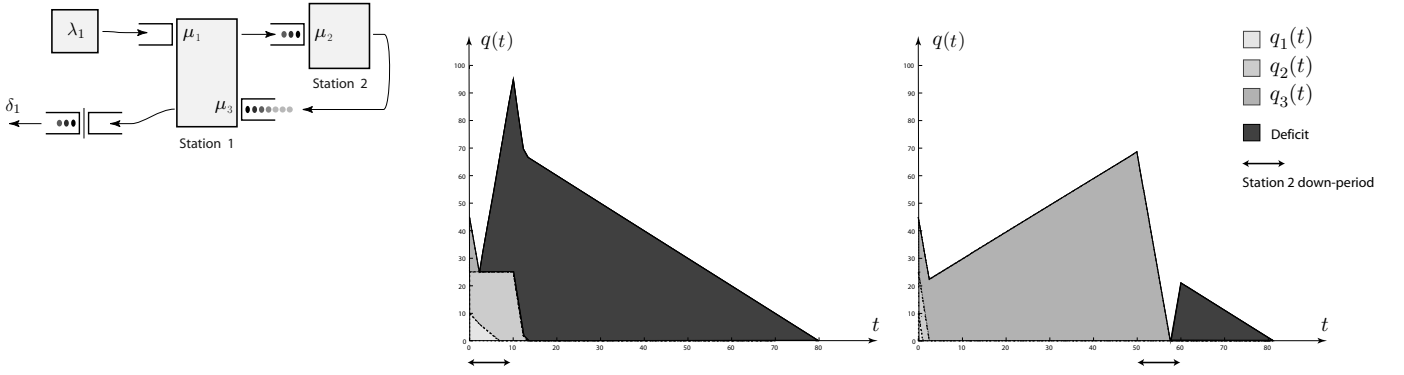


Fig. 7. The model considered is shown on the upper left-hand-side. The first plot shows buffer levels versus time for the GTO-B policy when resource 2 is inoperable for  $0 \leq t < 10$ . The plot at right shows buffer levels versus time for the GTO-M policy applied to the same network, with identical initial conditions, when resource 2 is inoperable for  $50 \leq t < 60$ . The impact of resource down-time is reduced significantly with advanced planning.

Given  $T^*(x; T_{\text{MTTR}}, t)$  for all  $x \in \mathbf{X}$  and  $t > 0$ , the greedy-time-optimal-with-breakdowns policy is defined as follows:

**GTO-B POLICY** Given  $0 \leq t < T_{\text{MTTR}}$  and  $q(t) = x$ , the allocation rate  $\zeta(t)$  is an optimizer of the linear program:

$$\begin{aligned}
 \min \quad & \eta \\
 \text{s. t.} \quad & \langle c^i, v \rangle \leq \eta, \quad \text{if } i \in \mathcal{I}_c(x), \\
 & \quad \quad \quad 1 \leq i \leq \ell_c, \\
 & \langle \nabla_x T^*(x), v \rangle = -1, \\
 & v_i \geq 0, \quad \text{if } x_i = 0, \\
 & \quad \quad \quad 1 \leq i \leq \ell, \\
 & v_i \leq 0, \quad \text{if } x_i = b_i, \\
 & \quad \quad \quad 1 \leq i \leq \ell, \\
 & B^{\text{down}} \zeta + \alpha = v, \\
 & \zeta \in \mathbf{U}^{\text{down}}, \quad v \in \mathbb{R}^\ell, \quad \eta \in \mathbb{R}.
 \end{aligned}$$

where  $T^*(x)$  denotes  $T^*(x; T_{\text{MTTR}}, t)$ .

The GTO-B policy empties the system in minimal time due to the dynamic programming condition  $\frac{d}{dt} T^*(q(t; x^0); T_{\text{MTTR}}, t) = -1$  for all  $0 \leq t \leq T^*(x^0; T_{\text{MTTR}}, 0)$ .

We consider simulations of the GTO-B for the pull model shown in the upper left-hand-side of Figure 7, with rates  $\delta_1 = 9$ ,  $\mu_1 = \mu_3 = 22$ , and  $\mu_2 = 10$ . The resulting vector load is  $\rho = (9/11, 9/10, 9/\mu_d)^T$ . The loss of either resource will cause significant disruption. A breakdown at station 2 is particularly significant due to its higher load.

In the simulation that follows, resource 2 is in-operable during  $[0, T_{\text{MTTR}}]$  with  $T_{\text{MTTR}} = 10$ , and the initial condition is  $x^0 = [10, 15, 20, 0, 0]^T$ . The plot on the left hand side of Figure 7 shows the buffer trajectories versus time under the GTO-B policy.

The inventory in  $q_2$  is quickly cleared once repair has been completed, after which the system works at capacity until the deficit created by the temporary loss of resource 2 has been cleared. Note that the unscheduled breakdown generated tremendous deficit. Assuming that the initial 20 units in buffer three were used to meet some of the demand, an additional 70 units of deficit accrued during repair.

### C. Preventative maintenance

We now examine the case where advance-warning is provided regarding system down-time. Our main conclusion is

that the impact of such disruptions is reduced significantly with advanced planning.

The time at which the resource is lost is denoted  $T_{\text{MTTF}}$ , and we continue to denote by  $T_{\text{MTTR}}$  the time required to bring the resource back into service. The acronym MTTF stands for *mean-time to failure*, which again reflects the fact that this time may not be known exactly in practice. However, for the purposes of control design, we again assume that this information is exact.

Given a time-period  $[T_{\text{MTTF}}, T_{\text{MTTF}} + T_{\text{MTTR}}]$  when a certain resource is not operational, the minimal draining time from the state  $q(t; x^0) = x$  is denoted  $T^*(x) = T^*(x; T_{\text{MTTF}}, T_{\text{MTTR}}, t)$ . This may be found through an obvious modification of (29). The greedy-time-optimal-with-maintenance policy is then defined as follows, based on the computation of  $T^*$ . We present the algorithm only for  $0 \leq t < T_{\text{MTTF}}$ . For  $t \in [T_{\text{MTTF}}, T_{\text{MTTF}} + T_{\text{MTTR}}]$  the GTO-B policy is used to define the allocation rate  $\zeta(t)$ .

**GTO-M POLICY** Given  $q(t) = x$ , with  $0 \leq t < T_{\text{MTTF}}$ , the allocation rate  $\zeta(t)$  is an optimizer of the linear program:

$$\begin{aligned}
 \min \quad & \eta \\
 \text{s. t.} \quad & \langle c^i, v \rangle \leq \eta, \quad \text{if } i \in \mathcal{I}_c(x), \\
 & \quad \quad \quad 1 \leq i \leq \ell_c, \\
 & \langle \nabla_x T^*(x), v \rangle = -1, \\
 & v_i \geq 0, \quad \text{if } x_i = 0, \\
 & \quad \quad \quad 1 \leq i \leq \ell, \\
 & v_i \leq 0, \quad \text{if } x_i = b_i, \\
 & \quad \quad \quad 1 \leq i \leq \ell, \\
 & B\zeta + \alpha = v, \\
 & \zeta \in \mathbf{U}, \quad v \in \mathbb{R}^\ell, \quad \eta \in \mathbb{R},
 \end{aligned}$$

where  $T^*(x)$  denotes  $T^*(x; T_{\text{MTTF}}, T_{\text{MTTR}}, t)$ .

As with the GTO-B policy, the GTO-M policy empties the system in minimal time.

As an illustration we return to the two-station pull-model considered in Section IV-B. We take  $T_{\text{MTTR}} = 10$  and  $T_{\text{MTTF}} = 50$ . The plot on the right hand side of Figure 7 shows the resulting buffer trajectory when the GTO-M policy is applied.

At the point where maintenance begins, there are about 70 units of inventory in  $q_3$  that is used to feed demand during the maintenance period. This surplus that has been staged at  $q_3$

is completely consumed by the demand by about time  $t = 57$ , and an additional 20 units of deficit is incurred before repair is complete at time  $t = 60$ .

#### D. Responding to persistent variability

So far, we have considered only large-scale, transient disturbances. In practice, there are also recurrent, persistent disturbances due to variability in demand and yield, and these must also be accounted for in the construction of an effective policy.

Suppose for example that a GTO policy is applied directly to the 16-buffer model in the presence of stochastic, persistent disturbances. The difficulty is most easily visualized by considering a low-dimensional workload relaxation. Consider the two-dimensional relaxation, with effective cost shown in Figure 3, and assume that  $(\mathbf{1} - \rho) \in \hat{W}^+$ . In this case the two dimensional fluid model admits a path-wise optimal solution with  $\mathcal{R}^*$  equal to the monotone region  $\hat{W}^+$ . If this region is used to construct a policy for the stochastic model with state process  $\mathbf{Q}$ , then one station will idle whenever  $\hat{W}(t; x^0) := \hat{\Xi}Q(t; x^0) \notin \hat{W}^+$ . Consequently, the workload process will exhibit significant chattering near the boundaries of  $\hat{W}^+$ , which will result in an excessively large mean value if variability is significant.

To illustrate how the region  $\mathcal{R}^*$  must be modified to account for variability we consider the following two dimensional Markov Decision Process (MDP) workload model with state process  $\mathbf{Y}$ , state space  $\mathbb{Z}^2$ , and one-step cost function equal to the effective cost on  $\mathbb{R}^2$  illustrated at left in Figure 3. The state process is defined in discrete-time via,

$$\mathbf{Y}(t+1) = \mathbf{Y}(t) + \mathbf{I}(t) + \mathbf{N}(t+1), \quad t \geq 0, \quad (30)$$

where  $\mathbf{I}$  is the two dimensional idleness process.

The process  $\mathbf{N}$  is i.i.d., and its marginal distribution is supported on  $\{(2, 0)^T, (0, 2)^T, (-5, -5)^T\}$ . Two cases were considered. In Case 1 the marginal distribution is uniform (each possible value occurs with probability  $1/3$ .) This results in a mean drift given by  $E[\mathbf{N}(t)] = -[1, 1]^T$ . In Case 2 the respective probabilities are given by  $\{43/240, 7/48, 31/120\}$ . The mean drift is given by  $E[\mathbf{N}(t)] = -[1, 1/10]^T$ , which is consistent with the drift vector shown in Figure 4. In each case, the second order statistics approximate the Central Limit Theorem variance obtained in a network with Poisson demand, and exponential servers.

The average-cost optimal policy is defined by a region  $\mathcal{R}^{\text{sto}} \subset \mathbb{Z}^2$  that determines the optimal idleness process  $\mathbf{I}^*$  as a function of  $\mathbf{Y}$ . Shown in Figure 8 are the optimal regions obtained using value iteration in each of the two cases. Note that the boundaries of  $\mathcal{R}^{\text{sto}}$  are approximately affine, as predicted by results in [6], [21].

Rather than compute the optimal policy for an MDP model, we propose here designs based on an affine enlargement of the cone  $\mathcal{R}^{\text{gto}}$  used for the fluid model. All of the policies considered in Sections 4.1-4.3 can be similarly modified.

We adopt the discrete-review structure described in [13] to define the following *Discrete-Review GTO policy*. We suppose that a time horizon  $T > 0$  is given, and the total allocation

$z(T)$  is obtained through a linear program. We then require that this total allocation is met by time  $T$ , but the details of the allocation  $\{z(t) : 0 \leq t \leq T\}$  are not specified. With this flexibility, it is possible to take into account the discrete nature of allocation decisions, and other issues such as set-up times that arise in manufacturing systems.

The affine enlargement of  $\mathcal{R}^{\text{gto}}$  is specified by a fixed target value  $\bar{x} \in \mathbb{R}_+^\ell$ . One goal in the GTO-DR policy introduced next is to maintain the lower bound  $Q(t) \geq \bar{x}$  for all  $t$ . For this purpose a parameter  $0 < \varepsilon_1 < 1$  is fixed and, given an initial condition  $x^0 \in \mathbf{X}$ , a surrogate target value is defined by

$$\bar{x}^1 = \bar{x}^1(x^0) := \min(\bar{x}, x^0 + \varepsilon_1 \bar{x}),$$

where the minimization is component-wise.

**GTO-DR ALGORITHM** Given the initial condition  $x^0$ ; the target state  $\bar{x} \in \mathbb{R}_+^\ell$ ; planning horizon  $T > 0$ ; and parameter  $\varepsilon_1 > 0$ , the GTO-DR control allocation is given by  $Z(T) = \zeta^{1*}T$ , where  $\zeta^{1*}$  is any optimizer to the following linear program:

$$\begin{aligned} \min \quad & \gamma \quad \text{s. t.} \quad \gamma \geq \langle c^i, y^1 \rangle, \quad i \in \mathcal{I}_c(y^1), \\ & y^1 = x^0 + (B\zeta^1 + \alpha)T_1, \\ & y^1 \geq \bar{x}^1, \\ & \bar{x} = y^1 + (B\zeta^2 + \alpha)T_2, \\ & \zeta^1, \zeta^2 \in \mathbf{U}, \quad y^1 \in \mathbf{X}, \quad \gamma \in \mathbb{R}, \end{aligned}$$

where  $T_1 := \min(T, T^*(x^0, \bar{x}))$ , and  $T_2 := T^*(x^0, \bar{x}) - T_1$ .

The solution  $\zeta^{1*}$  of the linear program in GTO-DR coincides with the allocation rate obtained in the GTO policy (27) when the target state  $\bar{x}$  is set to zero, and the planning horizon  $T$  is sufficiently small:

*Proposition 4.1:* Suppose that  $\bar{x} = \theta$  in the GTO-DR policy. Then, for each  $x^0 \in \mathbf{X}$ , there exists  $T > 0$  sufficiently small such that the solution  $\zeta^{1*} \in \mathbf{U}$  is a GTO allocation rate on  $[0, T]$  with initial condition  $x^0$ .

The GTO-DR policy is parameterized by the time horizon  $T$ , and the target-state  $\bar{x}$ . For a given value of  $T$  there are two major constraints to be taken into account on choosing the target state:

- (i) *Starvation avoidance:* For the fluid model, if  $q(0) \geq \bar{x}$ , then each resource  $1 \leq i \leq \ell_{ro}$  can work at capacity on the time-horizon  $[0, T]$ . That is, there exists  $v \in \mathbf{V}$ , satisfying

$$\langle \xi^i, v \rangle = -(1 - \rho_i), \quad (31)$$

for  $1 \leq i \leq \ell_{ro}$ , and  $\bar{x} + Tv \in \mathbf{X}$ . The following bound is suggested by results from [19], [2]: For some constant  $k_s > 0$ , and each  $1 \leq i \leq \ell_m$ ,

$$\sum_{j \geq 1} C_{i,j} \bar{x}_j \geq k_s \log\left(\frac{1}{1 - \rho}\right), \quad (32)$$

where  $C$  is the constituency matrix.

- (ii) *Appropriate hedging point:* The resulting affine region in workload space should be an appropriate enlargement of the fluid control region  $\mathcal{R}^{\text{gto}}$ : For a constant  $k_h > 0$ , and each  $1 \leq j \leq \ell_{ro}$ ,

$$\bar{w}_j := (\Xi \bar{x})_j \leq -k_h \left( \frac{1}{1 - \rho_j} \right). \quad (33)$$

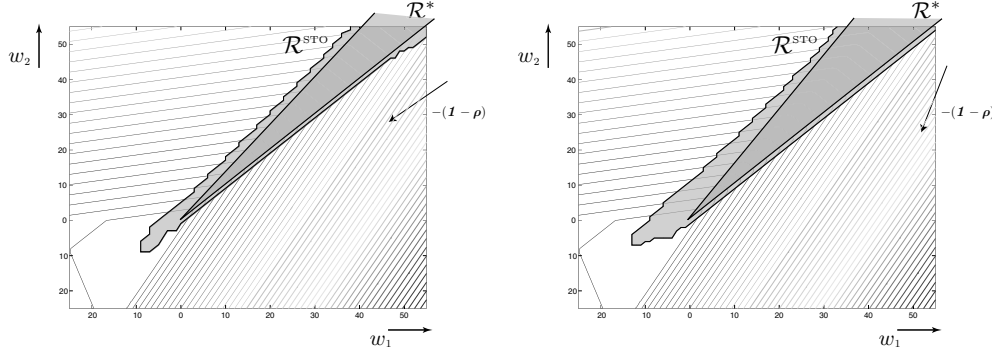


Fig. 8. Optimal policies for the two-dimensional workload model (30) of the production system shown in Figure 1. In each figure, the optimal stochastic control region  $\mathcal{R}^{\text{STO}}$  obtained using value iteration is compared with the optimal region  $\mathcal{R}^*$  obtained for the two dimensional fluid model.

The following specifications for the target value  $\bar{x}$  are based on consideration of the constraints (32) and (33) in the 16 buffer example: It will consist of  $O((1-\rho)^{-1})$  units of excess inventory (the two hedging points); and a total of  $O(\log((1-\rho)^{-1}))$  units of inventory at each station (interpreted as safety-stock values).

We conclude with a stability result for the stochastic model under the GTO-DR policy for a non-zero target state. Assumption (A4) formalizes the probabilistic assumptions imposed on (3). Under this condition, and with an appropriately chosen target value, the GTO-DR policy tracks the fluid idealization, and simultaneously ensures that critical resources do not risk starvation.

(A4) For all  $1 \leq i \leq \ell$  and  $1 \leq k \leq \ell_u$ , each of the stochastic processes  $\{A_i, R_{ik}, S_{ik}, t \geq 0\}$  is either null, or is an un-delayed renewal process whose increment process possesses a moment generating function that is bounded in a neighborhood of the origin.

In the following result we consider the state trajectory  $Q^\bullet$  under the GTO-DR policy. We find the policy is stabilizing, and nearly time-optimal. The result is based on consideration of the scaled processes, defined for  $r \geq 1$  via

$$q^{\bullet r}(t; x) = r^{-1} Q^\bullet(rt; rx), \quad x \in X, t \geq 0. \quad (34)$$

**Theorem 4.2:** Suppose that the network model (3) satisfies Assumptions (A1)-(A4) with  $X = \mathbb{R}_+^\ell$ . Assume moreover that for any fixed  $\bar{x} \in \mathbb{R}_+^\ell$ ,  $T > 0$ , there exists  $\varepsilon_1 > 0$  such that the constraint  $y^1 \geq \bar{x}^1$  in the GTO-DR policy is feasible for each initial condition. Then, there exists  $k_s^0 \geq 1$ , such that the following properties hold for the GTO-DR policy when the target value  $\bar{x}$  satisfies (31)–(33) with  $k_s \geq k_s^0$ , and any  $k_h \geq 0$ . The time-horizon is taken as  $T = \|\bar{x}\|$ .

(a) The controlled system is stable in the mean. This stability is uniform, in the sense that, for some  $b_c = b_c(k_s, k_h) < \infty$ , and any initial condition,

$$\limsup_{t \rightarrow \infty} E[c(Q^\bullet(t; x^0))] \leq b_c.$$

(b) Suppose that the GTO fluid trajectory  $q$  is unique for each initial condition. Then, the scaled process (34) *almost*

converges to  $q$ : For some  $b_0 < \infty$ ,  $I_0 > 0$ ,

$$\begin{aligned} \limsup_{r \rightarrow \infty} E_x[\|q^{\bullet r}(t; x^0) - q(t, x^0)\|] \\ \leq b_0 \exp(-I_0 k_s), \\ \text{where } x^0 \in X, \|x^0\| \leq 1, 0 \leq t \leq 2T^*(x^0). \end{aligned}$$

(c) The state trajectory is *almost* time-optimal for large initial conditions: There exists  $b_0 < \infty$ ,  $I_0 > 0$ , such that for any  $0 < \varepsilon \leq 1$  the GTO-DR policy satisfies

$$\begin{aligned} \limsup_{\|x^0\| \rightarrow \infty} P\{\|Q^\bullet(T^*(x^0); x^0)\| \geq \varepsilon \|x^0\|\} \\ \leq b_0 \exp(-I_0 k_s \varepsilon^2). \end{aligned}$$

(d) For any other admissible policy the scaled processes  $\{q^r\}$  satisfy for each  $x^0 \in \mathbb{R}_+^\ell$ , and  $0 \leq t \leq T^*(x^0)$ ,

$$\liminf_{r \rightarrow \infty} T^*(q^r(t; x^0)) \geq T^*(x^0) - t, \quad \text{a.s.}$$

**PROOF:** (sketch) The second result is similar to the main result in [16]. The idea is that the policy together with the statistical assumptions imply that the lower bound  $Q^\bullet(t) \geq \varepsilon_1 \bar{x}$  is met with high probability, of order  $1 - b_0 \exp(-I_0 k_s)$  for some constants  $b_0$  and  $I_0$ . It then follows that the constraints for the stochastic model and fluid model are virtually the same, so that resources for the stochastic model are non-idling with high probability when required by the GTO policy. Part (c) follows similarly using Large-Deviations estimates.

The uniform stability in (a) follows from (c) and [9, Theorem 4.1].

To establish (d), consider the set of all weak limits as  $r \rightarrow \infty$  of the stochastic processes  $\{q^r(t; x^0) := r^{-1} Q^\bullet(rt; rx^0) : r \geq 1, x^0 \in X\}$ . Any weak limit  $q$  is a solution to the fluid model equations (1) with initial condition  $x^0$ , and must satisfy the bound  $T^*(q(t)) \geq T^*(x^0) - t$  for all  $t \geq 0$ .  $\square$

## V. CONCLUSIONS

We have introduced in this paper several approaches to policy synthesis for complex, demand-driven networks. We have attempted to capture realistic constraints, as well as a range of control objectives. The model reduction techniques described in this paper are important in analysis, and critical in obtaining intuition regarding network behavior for the policies considered.

Of course, there is also much room for further research.

In our consideration of resource maintenance we assumed complete information regarding down-time. An obvious next step is to investigate the impact of uncertainty on performance for the algorithms described here, and to see if these algorithms can be improved given further statistical information.

It is natural to consider how these methods extend to models with competing players, and distributed information. An application of current interest concerns pricing and resource allocation in power distribution. Those who design deregulated markets can greatly benefit from a better understanding of the incentives in the network, which will eventually lead to a more efficient market design. Some preliminary results are contained in [7].

Finally, we look forward to testing these policies in a real-world setting. We believe that the approaches described here will have significant impact in network management for semiconductor and related manufacturing industries.

## REFERENCES

- [1] F. Avram, D. Bertsimas, and M. Ricard. Fluid models of sequencing problems in open queueing networks; an optimal control approach. In *Stochastic networks*, volume 71 of *IMA Vol. Math. Appl.*, pages 199–234. Springer, New York, 1995.
- [2] S.L. Bell and R.J. Williams. Dynamic scheduling of a system with two parallel servers: Asymptotic optimality of a continuous review threshold policy in heavy traffic. In *Proceedings of the 39th Conference on Decision and Control*, pages 1743–1748, Phoenix, Arizona, 1999.
- [3] M. Bramson. State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems: Theories and Applications*, 30(89-148), 1998.
- [4] H. Chen and A. Mandelbaum. Discrete flow networks: bottleneck analysis and fluid approximations. *Math. Oper. Res.*, 16(2):408–446, 1991.
- [5] H. Chen and David D. Yao. *Fundamentals of queueing networks: Performance, asymptotics, and optimization*. Springer-Verlag, New York, 2001. Stochastic Modelling and Applied Probability.
- [6] M. Chen, C. Pandit, and S.P. Meyn. In search of sensitivity in network optimization. *Queueing Systems*, 44(4):313–363, 2003.
- [7] I.-K. Cho and S. P. Meyn. The dynamics of the ancillary service prices in power distribution systems. In preparation. Preliminary version in the *Proceedings of the 42nd IEEE Conference on Decision and Control*, December 9-12, 2003.
- [8] A.J. Clark and H.E. Scarf. Optimal policies for a multi-echelon inventory problem. *Management Science*, Vol. 6, No. 4, 1960.
- [9] J.G. Dai and S.P. Meyn. Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control*, 40:1889–1904, 1995.
- [10] J.G. Dai and G. Weiss. A fluid heuristic for minimizing makespan in job shops. *Oper. Res.*, 50(4):692–707, 2002.
- [11] R. Dubrawski. Myopic and far-sighted strategies for control of demand-driven networks. Masters thesis, Department of Electrical Engineering, UIUC, 2000. Urbana, Illinois, USA.
- [12] S.B. Gershwin. *Manufacturing Systems Engineering*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [13] J.M. Harrison. The BIGSTEP approach to flow management in stochastic processing networks. In F. P. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, pages 57–90. Oxford Science Publications, Oxford U.K., 1996.
- [14] J.M. Harrison and J.A. Van Mieghem. Dynamic control of Brownian networks: state space collapse and equivalent workload formulations. *Ann. Appl. Probab.*, 7(3):747–771, 1997.
- [15] S.G. Henderson, S. P. Meyn, and V. Tadic. Performance evaluation and policy selection in multiclass networks. *Discrete Event Dynamic Systems: Theory and Applications*, 13:149–189, 2003. Special issue on learning and optimization methods (invited).
- [16] C. Maglaras. Dynamic scheduling in multiclass queueing networks: Stability under discrete-review policies. *Queueing Systems*, 31:171–206, 1999.
- [17] S.P. Meyn. Stability and optimization of queueing networks and their fluid models. In *Mathematics of stochastic manufacturing systems (Williamsburg, VA, 1996)*, pages 175–199. Amer. Math. Soc., Providence, RI, 1997.
- [18] S.P. Meyn. Sequencing and routing in multiclass queueing networks. Part I: Feedback regulation. *SIAM J. Control Optim.*, 40(3):741–776, 2001.
- [19] S. P. Meyn. Sequencing and routing in multiclass queueing networks. Part II: Workload relaxations. *SIAM J. Control Optim.*, 42(1):178–217, 2003.
- [20] S. P. Meyn. Dynamic safety-stocks for asymptotic optimality in stochastic networks. (submitted for publication), 2003.
- [21] S.P. Meyn. Value functions, optimization, and performance evaluation in stochastic network models. (submitted for publication), 2003.
- [22] J.K. Robinson, J.W. Fowler, and E. Neacy. Capacity loss factors in semiconductor manufacturing. Working paper, available from [www.FabTime.com](http://www.FabTime.com).
- [23] A. Sharifnia. Production control of a manufacturing system with multiple machine states. *IEEE Transactions on Automatic Control*, 33:600–626, July 1988.
- [24] G. Weiss. Optimal draining of a fluid re-entrant line. In Frank Kelly and Ruth Williams, editors, *Volume 71 of IMA volumes in Mathematics and its Applications*, pages 91–103, New York, 1995. Springer-Verlag.

PLACE  
PHOTO  
HERE

**Richard Dubrawski** was born in Pittsburgh, PA in 1967. He obtained a BSEE from The Pennsylvania State University in 1998 and an MSEE from UIUC in 2000. He is currently working at Viasat in San Diego as a senior Software Engineer.

PLACE  
PHOTO  
HERE

**Mike Chen** was born in Kaohsiung City, Taiwan in 1976. He obtained a BEng from McGill University in 1999 and an MS from UIUC in 2002, both in Electrical and Computer Engineering. He plans to complete a PhD in Electrical and Computer Engineering and an MS in Mathematics at UIUC in April, 2005.

PLACE  
PHOTO  
HERE

**Sean P. Meyn** received the B.A. degree in Mathematics Summa Cum Laude from UCLA in 1982, and the PhD degree in Electrical Engineering from McGill University in 1987. After a two year postdoctoral fellowship at the Australian National University in Canberra, Dr. Meyn and his family moved to the Midwest. He is now a Professor at the Department of Electrical and Computer Engineering, and a Research Professor at the Coordinated Science Laboratory at the University of Illinois.

Dr. Meyn has served on the editorial boards of several journals in the systems and control, and applied probability areas. He is coauthor with Richard L. Tweedie of the monograph *Markov Chains and Stochastic Stability*, Springer-Verlag, London, 1993; and received jointly with Tweedie the 1994 *ORSA/TIMS Best Publication In Applied Probability Award*.