SEQUENCING AND ROUTING IN MULTICLASS QUEUEING NETWORKS PART I: FEEDBACK REGULATION

SEAN P. MEYN*

Abstract.
This paper establishes new criteria for stability and for instability of multiclass network models under a given stationary policy. It also extends previous results on the approximation of the solution to the average cost optimality equations through an associated fluid model: It is shown that an optimized network possesses a fluid limit model which is itself optimal with respect to a total cost criterion.

A general framework for constructing control algorithms for multiclass queueing networks is proposed based on these general results. Network sequencing and routing problems are considered as special cases. The following aspects of the resulting *feedback regulation policies* are developed in the paper:

(i) The policies are stabilizing, and are in fact geometrically ergodic for a Markovian model.

(ii) Numerical examples are given. In each case it is shown that the feedback regulation policy closely resembles the average-cost optimal policy.

(iii) A method is proposed for reducing variance in simulation for a network controlled using a feedback regulation policy.

Key words. Queueing Networks, Routing, Scheduling, Optimal Control

AMS subject classifications. Primary 90B35, 68M20, 90B15; Secondary 93E20, 60J20.

1. Introduction. This paper concerns the effective management of large networks through scheduling and routing.

Specific applications of interest include cellular and internet communication systems, large scale manufacturing processes, and computer systems (see e.g. [5, 55, 22]). In spite of the diversity of these applications, one can find many common goals:

(i) Controlling delay, throughput, inventory, and loss. The crudest issue is *stability*: do queue lengths remain bounded for all time?

(ii) Estimating performance, or comparing the performance of one policy over another one. *Performance* is context-dependent, but common metrics are average delay, and loss probabilities.

(iii) Prescriptive approaches to policy synthesis which are intuitive, flexible, robust, and reasonable in complexity. *Robustness* means that the policy will be effective even under significant modeling error. By *flexibility* we mean that the policies will react appropriately to changes in network topology, or other gross structural changes.

(iv) In applications to telecommunications or power systems one may have limited information. Routing or sequencing decisions must then be determined using only that information which can be made available. This issue is becoming less critical with ever-increasing information processing power. In the future internet it may be possible to assume essentially complete information at every node in the network through current flooding algorithms, and proposed *explicit congestion notification algorithms* [21].

There are currently several popular classes of models which describe in various levels of detail the dynamics of a queueing network to address the range of issues in

*Coordinated Science Laboratory and the University of Illinois, 1308 W. Main Street, Urbana, IL 61801, URL `http://black.csl.uiuc.edu:80/~meyn` (s-meyn@uiuc.edu). Work supported in part by NSF Grants ECS 9403742 and ECS 99 72957

(i)-(iv). The utility of a particular model depends upon ones particular goals.

A traditional academic approach to policy synthesis is to construct an MDP (Markov decision process) model for the network. This involves constructing a controlled transition operator $P_a(x, y)$, which gives the probability of moving from state x to state y when the control decision $a \in \mathcal{A}(x)$ is applied. The state space X (where x and y live) is typically taken as the set of all possible buffer levels at the various stations in the network; $\mathcal{A}(x)$ is then the set of feasible control actions when the state takes the value $x \in \mathsf{X}$. Given an MDP model, and a one step cost function $c\colon \mathsf{X} \to \mathbb{R}_+$, a solution to the average cost optimal control problem is found by solving the resulting dynamic programming equations,

$$\eta^* + h^*(x) = \min_{a \in \mathcal{A}(x)} [c(x) + P_a h^*\,(x)] \tag{1.1}$$

$$F^*(x) = \arg\min_{a \in \mathcal{A}(x)} P_a h^*\,(x), \qquad x \in \mathsf{X}. \tag{1.2}$$

The function F^* on X then defines an optimal stationary policy.

The difficulty with this approach is very well known: When buffers are infinite, this becomes an infinite dimensional optimization problem. Even when considering finite buffers, the complexity grows exponentially with the dimension of the state space. Some form of aggregation is necessary - the Markovian model is simply too detailed to be useful in optimization.

An elegant approach is to consider the model in heavy traffic where a reflected Brownian motion model is appropriate. The papers [24, 28] develop these ideas for the network scheduling or sequencing problems, and [32] considers routing and other control problems. One is then faced with optimizing a controlled stochastic differential equation (SDE) model. In many examples considered in the literature this control problem has a simple intuitive solution. This is just one example of a fluid model for the physical network. Another popular model is the "σ-ρ" constrained fluid model [11, 51], and the linear fluid model considered here (see e.g. [8, 7, 56, 41, 42, 2]). Any one of these models is valuable in network design because unimportant details are stripped away.

Justification for considering these various idealizations comes from theory that establishes solidarity between idealized fluid models, and more accurate discrete models, when the load is close to capacity [32, 4], or the state of the system is large (e.g. the network is congested [45], or a 'large deviation' occurs [54]). Stability theory for networks, as in (i), has essentially reached maturation over the past decade, following counter-examples introduced in [35, 52]. This theory is based on the close ties between a stochastic network model, and its linear fluid counterpart [14, 15, 16].

There are, however, several difficulties with these approaches:

- The Brownian motion approximation is based on a model in heavy traffic. If the stations are not balanced then one loses some information at the stations which are not heavily loaded;
- Although the optimal control problem for a Brownian motion or σ-ρ constrained fluid model often has a simple intuitive solution, frequently this is not the case. There is currently no general practical method for generating policies;
- It is not always obvious how to translate an optimal policy for an abstract model to a feasible and efficient policy for the original discrete model.

In this paper, we consider exclusively the linear fluid model (2.2) in design. Theorem 3 below, an extension of the main result of [45], establishes a strong form of solidarity between the discrete optimization problem (1.1,1.2), and a related total-cost optimal control problem for the linear fluid model. These results can be generalized to show that an optimized SDE model possesses a fluid limit model which is itself optimal with respect to the total-cost criterion. Hence to optimize the Brownian motion model one must also solve the linear fluid model optimization problem.

Translation of a policy from the fluid model to the original discrete model of interest is again not obvious. This issue is addressed in [42, 2] where it is shown that virtually any 'fluid trajectory' can be approximately tracked using a discrete policy for the discrete-stochastic network. In [13, 45] the results from several numerical studies are described. It is found that the optimal discrete policy resembles the optimal fluid policy, but with the origin θ for the model (2.2) shifted to some value $\bar{x} \in \mathbb{R}^{\ell}_{+}$. From these results it is argued that one should use the fluid model to attempt to regulate the state to some deterministic value $\bar{x}$. In the numerical studies considered, it was found that the average cost was nearly optimal, and that the variance at each buffer was *lower* than that obtained using an optimal policy. Computation of optimal policies for the linear fluid model appears to be feasible for network models of moderate complexity [56, 41, 49].

A final motivation for considering the simplest network model follows on considering our main goal: robust policy synthesis. As described in (iii) above, any policy that we construct should be sufficiently robust so that it will tolerate modeling error resulting from uncertain variability in service or arrival rates.

The main results of the present paper builds upon those of [45, 44]:

(i) The class of models is extended to include routing and processor sharing models, as well as the scheduling models considered earlier. This requires that we introduce a notion of stabilizability for networks, which is a generalization of the usual capacity conditions.

(ii) The underlying theory is improved. It is shown that optimal policies have optimal fluid limits with respect to the total cost criterion. This improves upon the main result of [45, 44] which required specialization to a class of 'fluid limit models' obtained via weak convergence. Moreover, the stability theory relating networks and their fluid models is improved to give criteria for geometric ergodicity, and simpler conditions implying transience of the controlled network.

(iii) The practical usefulness of the approach is improved by borrowing from the BIGSTEP approach of [28], and by exploring reduced complexity control approaches for the fluid model.

(iv) Numerical examples are given. In each case it is shown that the feedback regulation policy closely resembles the average-cost optimal policy.

(v) Consideration of the fluid model leads to an approach to estimating steady state performance indicators, such as mean delay, through simulation. This requires care since standard Monte Carlo simulation is known to have high variance for highly loaded networks.

The viewpoint arrived at in this paper leads to policies which are similar to those found through a heavy traffic analysis using a Brownian motion approximation. Consider for example the models treated in [32]. In each case one could perform designs on the fluid model, translate these policies as in (4.1), and arrive at the same policy that was obtained using a Brownian motion approximation. Given the greater complexity of the Brownian motion model, we conclude that while diffusion approximations are tremendously useful for analysis and performance approximation, presently they appear to be less useful for the purposes of control design. This will

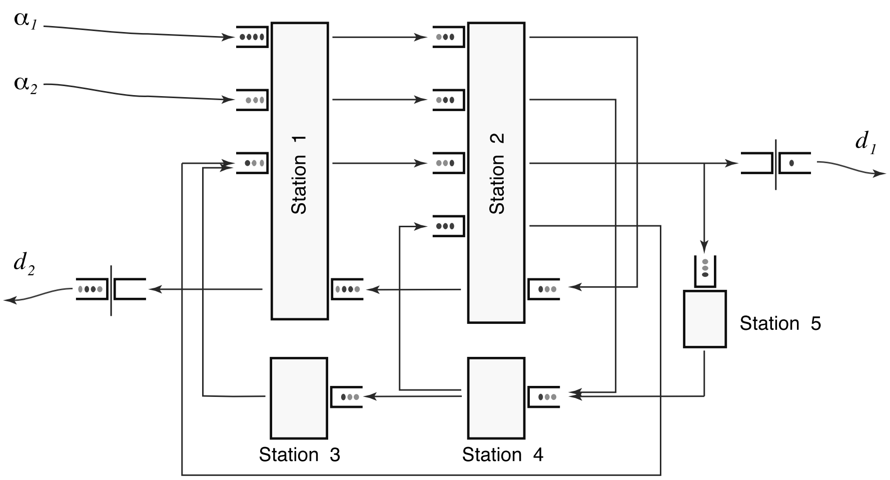

FIG. 1. *A network with many buffers, controlled routing, uncontrolled routing, multiple demands, and virtual buffers.*

change if more efficient numerical methods can be devised for control synthesis in SDE models [38].

One of the most important benefits of a heavy traffic assumption is that the resulting "state space collapse" can result in a model of reduced complexity. In the models considered in the aforementioned references, in each case one is left with a one dimensional state process which captures all relevant information. This model reduction is obtained for *either* model, fluid or SDE, when the system load approaches a critical value. The aim of Part II is to exploit this observation to prove that, under certain geometric conditions, an optimal policy for the fluid model may be translated to form a policy which is approximately optimal for the stochastic model [43].

The remainder of the paper is organized as follows. Section 2 describes the class of models considered, and their associated fluid limit model. Here some general stability theory for networks and their fluid models is presented, including criteria for geometric ergodicity. In Section 3 this solidarity between the fluid model and the discrete network is extended. It is shown that, provided the fluid model is stabilizable, there exists an average cost optimal policy whose fluid model is optimal with respect to the total cost criterion. Several examples are given to illustrate the relationship between the two optimization problems for specific models. The *feedback regulation* policies are introduced in Section 4. Several classes of stabilizing fluid policies are described, and a stability proof is provided in this section. Conclusions are postponed to Part II.

2. Networks and their fluid limit models. The networks envisioned here consist of a finite set of resources, a finite set of buffers, and various customers classes which arrive to the network for processing. Resources perform various activities, which transform parts or customers at the various buffers. On completion of service, a customer either leaves the network, or visits another resource for further processing. This is the intuitive definition of a multiclass queueing network. A popular continuous-time model is given by,

$$Q(t;x) = x - S(Z(t;x)) + R(Z(t;x)) + A(t), \qquad t \geq 0. \tag{2.1}$$

The vector-valued stochastic process $Q(t;x)$ denotes the buffer levels at time t, with initial condition $Q(0;x) = x \in \mathbb{R}^\ell$. Some of these buffers may be *virtual*. In a manufacturing model, such as that shown in Figure 1, virtual buffers may correspond to backlog or excess inventory.

The vector-valued stochastic process $Z(t;x)$ is the *allocation* (or *control*). The ith component $Z_i(t;x)$ gives the cumulative time that the activity i has run up to time t.

The vector-valued process $\boldsymbol{A}$ may denote a combination of exogenous arrivals to the network, and exogenous demands for materials *from* the network. The vector-valued functions $S(\,\cdot\,), R(\,\cdot\,)$ represent, respectively, the effects of random service rates, and the effects of a combination of possibly uncontrolled, possibly random routing, and random service rates.

The *fluid limit model* is obtained on considering a congested network. When $Q(t;x)$ is large in magnitude, variations in the arrival and service processes appear small when compared with the state. The behavior of $\boldsymbol{Q}$ when viewed on this large spatial scale will appear deterministic, and can be approximated by the *mean field equations*, or *(linear) fluid model*,

$$q(t;x) = x + Bz(t;x) + \alpha t, \qquad t \geq 0. \tag{2.2}$$

Here B is a matrix of appropriate dimension, interpreted as a long-run average of $\boldsymbol{R} - \boldsymbol{S}$, and α is a long-run average of $\boldsymbol{A}$.

There are several ways of making this precise, and very few assumptions are required to ensure the existence of a well defined fluid limit model. A construction is provided in Section 2.1, and Section 2.2 describes stability theory for (2.1) based on the simpler model (2.2).

Section 2.1 introduces a discrete-time countable state-space MDP model. In the special case where all of the driving processes $(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{S})$ are multidimensional Poisson processes, the discrete-time model is obtained from (2.1) via *uniformization* [40]. The MDP model is convenient for the purposes of optimization, and also provides the simplest setting for constructing the fluid limit model through scaling $\boldsymbol{Q}$ and $\boldsymbol{Z}$.

2.1. A Markovian network model and its fluid limit. Consider the M/M/1 queue, described in continuous time by

$$Q(t;x) = x - S(Z(t;x)) + A(t), \qquad t \geq 0,$$

where the *cumulative busy time* $\boldsymbol{Z}$ satisfies $\frac{d}{dt}Z(t;x) = 1$ whenever $Q(t;x) \neq 0$. The stochastic processes $(\boldsymbol{A}, \boldsymbol{S})$ are one-dimensional Poisson processes with rates α, μ respectively. The fluid model is given by the analogous equation,

$$q(t;x) = -\mu z(t) + \alpha t, \qquad t \geq 0. \tag{2.3}$$

To obtain a discrete-time process, one might sample at successive jump times of $\boldsymbol{Q}$, but this would introduce bias. When $Q(t;x) = 0$, only upward jumps are possible, hence sampling is less frequent in this situation, and the overall sampling rate is *non-uniform*. Consequently, the steady-state queue length for the sampled process is strictly larger than that of the unsampled queue.

Uniformization corrects this by introducing *virtual service times*, and sampling $\boldsymbol{Q}$ at times of arrivals, service completions, or virtual service completions. For example, the first sampling time is of the form $\tau_1 = \min(S_1, T_1)$, where S_1 and T_1 are exponentially distributed random variables with mean $1/\mu$ and $1/\alpha$, respectively. If

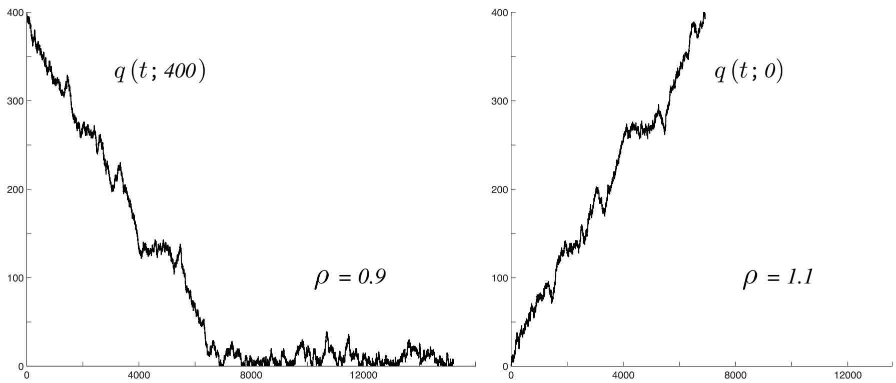

FIG. 2. *The M/M/1 queue: In the stable case on the left we see that the process $Q(t;x)$ appears piecewise linear, with a relatively small high frequency 'disturbance'. The process explodes linearly in the unstable case shown at right.*

$Q(0;x) = x > 0$, then S_1 is the remaining service time for the customer currently in service. If $x = 0$, then S_1 is again exponential with mean $1/\mu$, but it is now a random variable which is independent of the original queue length process. Sampling results in a discrete-time Markov chain with transition probabilities,

$$P(x, x+1) = \alpha, \qquad P(x, (x-1)_+) = \mu, \qquad x = 0, 1, 2, 3, ...$$

When $\rho = \alpha/\mu$ is less than one, then the queue is *positive recurrent*, as shown in the left hand side of Figure 2.

The discrete-time M/M/1 queue model may be viewed as a random linear system,

$$Q(k+1) = Q(k) + \widetilde{B}(k+1)U(k) + \widetilde{\alpha}(k+1), \tag{2.4}$$

where the sequence $\boldsymbol{U}$ is defined again by the non-idling policy $U(k) = \mathbb{I}(Q(k) > 0)$, $k \geq 0$, and $\{\widetilde{B}(k), \widetilde{\alpha}(k) : k \geq 1\}$ is an i.i.d. sequence satisfying,

$$\begin{pmatrix} \widetilde{B}(k) \\ \widetilde{\alpha}(k) \end{pmatrix} = \begin{cases} -e^1 & \text{with prob. } \mu \\ e^2 & \text{with prob. } \alpha, \end{cases}$$

with e^i equal to the standard basis element in $\mathbb{R}^2$, $i = 1, 2$. The general network model may be sampled in the same way to obtain a similar, multidimensional model whenever $(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{S})$ are Poisson.

Since we will not consider again the continuous time model, we will denote by $(\boldsymbol{Q}, \boldsymbol{U}) = \{(Q(k;x), U(k;x)) : k \geq 0\}$ the state-allocation process for a discrete-time model with initial condition x (this dependency will be supressed when the particular initial condition is not relevant). We assume that there are ℓ buffers, and ℓ_u activities, so that $(\boldsymbol{Q}, \boldsymbol{U})$ evolves on $\mathbb{Z}_+^\ell \times \mathbb{Z}_+^{\ell_u}$. As in the continuous-time case, each $U_i(k)$, $1 \leq i \leq \ell_u$, takes binary values. The discrete-time Markovian model is then *defined* as the random linear system (2.4).

We assume that the sequence $\{\widetilde{B}(k), \widetilde{\alpha}(k) : k \geq 1\}$ is i.i.d.. For each k, the matrix $\widetilde{B}(k)$ has dimension $\ell \times \ell_u$, the vector $\widetilde{\alpha}(k)$ has dimension ℓ, and the components of both take on integer values, again strictly bounded. We assume moreover that the components of $\widetilde{\alpha}(k)$ are non-negative.

This is a generalization of the sampled model, in which the entries of $(\widetilde{B}(k), \widetilde{\alpha}(k))$ take on the values $(-1, 0, 1)$ only. However, this discrete-time model covers only a very narrow set of stochastic network models. For example, it is not possible to convert the natural continuous time model into a countable-state, discrete-time MDP if service times are uniformly distributed. We restrict ourselves to the simple discrete-time model for the sake of exposition only. General distributions are considered in [14, 15] where it is shown that stability theory goes through without change. To generalize the results of the present paper, e.g. Theorem 4, one must assume a bounded hazard rate as in [46] to ensure that the mean forward recurrence time is bounded. Part II, which does not require a Markovian description, develops the general model (2.1) [43].

We assume that there is an integer $\ell_m \geq 1$, and an $\ell_m \times \ell_u$ *constituency matrix* C, such that

$$CU(k) \leq \mathbf{1}, \qquad k \geq 0,$$

where $\mathbf{1}$ denotes a vector of ones. The entries of C take on binary values, and each row defines a *resource*: The ith resource $\mathcal{R}_i$ is defined to be the set of activities j such that $C_{ij} = 1$.

There may also be auxiliary constraints on the control sequence $\boldsymbol{U}$, and further constraints on $\boldsymbol{Q}$. For example, buffers may require synchronous processing, or strict limits on buffer levels may be imposed. We assume that these may be expressed through linear constraints,

$$C_a U(k) \leq b_a, \qquad C_s Q(k) \leq b_s, \qquad k \geq 0,$$

for matrices C_a, C_s, and vectors b_a, b_s of appropriate dimension.

We have thus restricted $(\boldsymbol{Q}, \boldsymbol{U})$ to lie in the countable sets,

$$Q(k) \in \mathsf{X} \cap \mathbb{Z}^{\ell} \quad U(k) \in \mathsf{U} \cap \mathbb{Z}^{\ell_u}, \qquad k \geq 0,$$

where

$$\mathsf{U} := \{\zeta \in \mathbb{R}^{\ell_u}_+ : \zeta \geq \theta,\ C\zeta \leq \mathbf{1},\ C_a\zeta \leq b_a\}\,; \tag{2.5}$$

$$\mathsf{X} := \{x \in \mathbb{R}^{\ell}_+ : x \geq \theta,\ C_s x \leq b_s\}. \tag{2.6}$$

We assume throughout the paper that U is bounded. Unless noted otherwise, we assume that C_a and C_s are null.

The sequence $\boldsymbol{U}$ is an adapted (history dependent) stochastic process. We say that $\boldsymbol{U}$ is defined by a *stationary policy* if there is a feedback function $F \colon \mathsf{X} \to \mathsf{U}$ satisfying

$$\mathsf{P}(U_i(k) = 1 \mid Q(0), \dots, Q(k)) = F_i(Q(k)), \qquad k \geq 0.$$

The policies we consider are primarily stationary, or based on such policies.

In the classical *scheduling model* the ℓ_m resources $\{\mathcal{R}_i : 1 \leq i \leq \ell_m\}$ are called *stations*. There is one activity for each customer class, giving $\ell_u = \ell$. Class i customers wait in the ith queue, if necessary, and then receive service via the ith activity. Upon completion of service, a class i customer becomes a class j customers with probability R_{ij}, and exits the system with probability $R_{i0}{:=}1-\sum_j R_{ij}$. The constraint $CU(k) \leq \mathbf{1}$ is the usual condition that no two customers receive service simultaneously at a single station.

To construct the fluid limit model, first consider the time-invariant means, given by

$$B = \mathsf{E}[\widetilde{B}(k)], \quad \alpha = \mathsf{E}[\widetilde{\alpha}(k)], \qquad k \geq 1.$$

We may then write

$$Q(k+1) = Q(k) + BU(k) + \alpha + D(k+1), \tag{2.7}$$

where the process $\boldsymbol{D}$ is bounded, and it is a martingale difference sequence with respect to the natural filtration. In this way, the model (2.4) may be viewed as a deterministic 'fluid model', with a bounded 'disturbance' $\boldsymbol{D}$. When the initial condition $Q(0)$ is large, then the state dominates this disturbance, and the network behavior appears deterministic.

The fluid limit model considered in this paper is obtained by scaling the process, both temporally and spatially, through a scaling parameter $n \in \mathbb{Z}_+$. For any $x \in \mathbb{R}_+^\ell$, $n \geq 1$, we define,

$$q^n(t;x) = \frac{1}{n} Q(nt;nx)\,; \tag{2.8}$$

$$z^n(t;x) = \frac{1}{n} \sum_{i \leq nt} U(i;x), \qquad t = \tfrac{0}{n}, \tfrac{1}{n}, \tfrac{2}{n}, \ldots \tag{2.9}$$

where we are taking the integer part of nx whenever necessary so that the initial condition nx lies in $\mathbb{Z}_+^\ell$. We then extend the definition of $\{q^n(t;x), z^n(t;x)\}$ to arbitrary $t \in \mathbb{R}_+$ so that these processes are linear on each time segment $[i/n, (i+1)/n]$, and continuous on $\mathbb{R}_+$.

We have $q^n(0;x) = x$, $z^n(0;x) = \theta$, and $\{q^n(\,\cdot\,;x), z^n(\,\cdot\,;x)\}$ are Lipschitz continuous for any n and x. Typically, we find that $\{\boldsymbol{q}^n, \boldsymbol{z}^n\}$ converges to a limiting, deterministic function of time $\{\boldsymbol{q}, \boldsymbol{z}\}$ as $n \to \infty$. The limits $\boldsymbol{q}$ and $\boldsymbol{z}$ are piecewise linear functions of t in all of the examples considered below. Figure 2 illustrates the nature of this convergence for the M/M/1 queue, where the limiting process satisfies (2.3).

For each x, the set $\mathcal{L}_x$ denotes all weak limits of $\{(q^n(\,\cdot\,;x), z^n(\,\cdot\,;x)) : n \geq 1\}$ as $n \to \infty$. The *fluid limit model*, denoted $\mathcal{L}$, is the union of $\mathcal{L}_x$ over all initial conditions. Any $(q,z) \in \mathcal{L}$ satisfies equations (2.2), together with the rate-constraints,

$$C[z(t) - z(s)] \leq (t-s)\mathbf{1}, \quad z(t) - z(s) \geq \theta, \qquad t \geq s \geq 0. \tag{2.10}$$

That is, $\frac{z(t)-z(s)}{t-s} \in \mathsf{U}$ for any $t \neq s$.

2.2. Stability of the models. Stability of the network under some policy requires some assumptions on the model. We say that $\boldsymbol{z}$ is a *feasible allocation* for the fluid model if the resulting state trajectory $\boldsymbol{q}$ satisfying (2.2), (2.10) remains in X for all $t \geq 0$. The fluid model is said to be

(i) *Stabilizable* if, from any initial condition $x \in \mathsf{X}$, there exists $T_\theta < \infty$ and a feasible allocation $\boldsymbol{z}$ such that

$$q(t;x) = x + \alpha t + Bz(t) = \theta, \qquad t \geq T_\theta.$$

(ii) *Controllable* if for any pair $x, y \in \mathsf{X}$, there is a feasible allocation $\boldsymbol{z}$, and a time T, such that $q(T;x) = y$.

Note that one can assume without loss of generality that an allocation $\boldsymbol{z}$ driving x to y is *linear*:

PROPOSITION 1. *Suppose that $x, y \in \mathsf{X}$, $T > 0$, and $\boldsymbol{z}$ is an allocation satisfying*

$$q(T;x) = x + \alpha T + Bz(T) = y.$$

Then the linear allocation $z^1(t) = \bar{z}t$, $0 \le t \le T$, also brings $\boldsymbol{q}$ to y from x at time T, and satisfies (2.10) on $[0,T]$. □

A necessary condition for stabilizability is that there exist some solution to the equilibrium equation

$$B\zeta^{\mathrm{ss}} + \alpha = \theta, \qquad \zeta^{\mathrm{ss}} \in \mathsf{U}. \tag{2.11}$$

In the special case of network scheduling, the $\ell \times \ell$ matrix B has the form

$$B = -(I - R^{\mathrm{T}})M, \tag{2.12}$$

where M is the diagonal matrix with diagonal entries $\mu^{\mathrm{T}} = (\mu_1, \dots, \mu_\ell)$. There is a unique solution to the equilibrium equation (2.11), given by $\zeta^{\mathrm{ss}} = -B^{-1}\alpha$, and the standard load condition may be written,

$$\vec{\rho} = -C\zeta^{\mathrm{ss}} = M^{-1}(I - R^{\mathrm{T}})^{-1}\alpha < \mathbf{1}. \tag{2.13}$$

It is readily seen that the load condition implies stabilizability. In routing models and many other examples the 'load' at a station is policy dependent [32].

To obtain sufficient conditions for stabilizability, it is convenient to envision (2.2) as a differential inclusion,

$$\dot{q} \in \mathsf{V} := \{B\zeta + \alpha : \zeta \in \mathsf{U}\} \subset \mathbb{R}^\ell.$$

The set V is equal is the set of possible velocity vectors for the fluid model. We let $-\mathsf{V}$ denote its reflection. The proof of the following result is obvious and will be omitted. In Proposition 2 the set $B(\theta,\varepsilon)$ denotes the open ball of radius ε, centered at the origin.

PROPOSITION 2. *The fluid model (2.2) is*

(i) *stabilizable if and only if there exists $\varepsilon > 0$ such that*

$$B(\theta,\varepsilon) \cap \mathbb{R}^\ell_+ \subset \{-\mathsf{V}\} \cap \mathbb{R}^\ell_+$$

(ii) *controllable if and only if there exists $\varepsilon > 0$ such that $B(\theta,\varepsilon) \subset \mathsf{V}$.*

□

Either of the conditions (i) or (ii) can be formulated as a finite linear program. For example, the following set of constraints summarizes the condition that V contains each of the vectors $\{-\varepsilon e^i : 1 \le i \le \ell\}$.

$$\begin{aligned} B\zeta^i + \alpha &= -\varepsilon_i e^i \\ C\zeta^i &\le \mathbf{1} \\ \zeta^i &\ge \theta, \qquad 1 \le i \le \ell. \end{aligned}$$

A related linear program is devised to define the system load in [26, 25].

We now turn to the discrete-stochastic network.

When controlled by a randomized stationary policy with feedback law F, the state process becomes a time-homogeneous Markov chain. The state transition matrix is denoted P_F - the subscript is supressed when there is no risk of ambiguity. *Stability* of the controlled process is defined as positive recurrence of the resulting Markov chain. Under stability, there exists a unique invariant probability $\pi = \pi_F$, and steady state performance measures such as mean delay, or average total congestion can be described in terms of the invariant probability.

We assume that, under the transition law P, the state process possesses a single communicating class $\mathcal{C}$ which contains the origin θ. We assume moreover that the controlled system is ψ-irreducible and aperiodic, as defined in [47]. Defining the first return time to a set $A \subseteq \mathsf{X}$ by

$$\tau_A = \min(k \geq 1 : Q(k) \in A),$$

the ψ-irreducibility condition can be expressed $\mathsf{P}_x(\tau_\theta < \infty) > 0$, for any $x \in \mathsf{X}$. This is typically a minor constraint on the policy. For the network scheduling problem these conditions hold when the policy is non-idling.

Throughout the paper we use $c\colon \mathbb{R}_+^\ell \to \mathbb{R}_+$ to denote a norm, i.e. it is continuous, convex, vanishes only at θ, and it is radially homogeneous. The function c will be interpreted as a one-step cost function for the model. For a particular stationary policy, the controlled chain is called *c-regular* if for any initial condition x,

$$\mathsf{E}_x\Big[\sum_{i=0}^{\tau_\theta - 1} c(Q(i))\Big] < \infty.$$

A c-regular chain always possesses a unique invariant probability π such that

$$\pi(c) := \sum_{x \in \mathsf{X}} c(x)\, \pi(x) < \infty.$$

A stationary Markov policy (and its associated feedback function F) is called *regular* if the controlled chain is c-regular. In this case it follows from the f-norm ergodic theorem of [47, Chapter 14] that the following average cost exists and is independent of the initial condition x:

$$\begin{array}{llcl} \textbf{(i)} & J(x,w) = \lim_{k\to\infty} \mathsf{E}_x[c(Q(k))] & = & \pi(c). \\ \textbf{(ii)} & \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} c(Q(k)) & = & \pi(c), \qquad a.s. \end{array}$$

The fluid limit model is said to be *stable* if there exists $\varepsilon > 0$ and $T < \infty$ such that $q(T; x) = \theta$ for any $q \in \mathcal{L}_x$ with $\|x\| \leq \varepsilon$. It will be called *L_p-stable* if for some $\varepsilon > 0$,

$$\lim_{t\to\infty} \sup_{q \in \mathcal{L}_x : \|x\| \leq \varepsilon} \mathsf{E}[\|q(t)\|^p] = 0.$$

The following result is a minor generalization of [45, Theorem 5.2]. Related results are found in [20, 52, 14, 15].

Theorem 3. *The following stability criteria are equivalent for the network under any nonidling, stationary Markov policy.*

(i) *There exists $b_0 < \infty$, and a function $V\colon \mathbb{R}^\ell \to \mathbb{R}_+$ such that the drift condition holds,*

$$PV\,(x) := \mathsf{E}_x[V(Q(k+1)) \mid Q(k) = x] \leq V(x) - c(x) + b_0, \qquad x \in \mathsf{X}. \tag{2.14}$$

The function V is equivalent to a quadratic in the sense that, for some $\varepsilon > 0$,

$$1+\varepsilon\|x\|^2 \le V(x) \le 1+\varepsilon^{-1}\|x\|^2, \qquad x\in \mathsf{X}.$$

(ii) *For some quadratic function V and some $b_0<\infty$,*

$$\mathsf{E}_x\Bigl[\sum_{n=0}^{\tau_\theta} c(Q(n))\Bigr] \le V(x)+b_0, \qquad x\in\mathsf{X}.$$

(iii) *For some quadratic function V and some $b_0<\infty$,*

$$\sum_{n=1}^{N}\mathsf{E}_x[c(Q(n))] \le V(x)+b_0N, \quad \textit{for all } x \textit{ and } N\ge 1.$$

(iv) *The fluid limit model is L_2-stable.*

(v) *The total cost for the fluid limit is uniformly bounded in the sense that, for some quadratic function V,*

$$\mathsf{E}\Bigl[\int_0^\infty \|q(\tau;x)\|\,d\tau\Bigr] \le V(x), \qquad x\in\mathbb{R}^\ell_+,\ q\in\mathcal{L}_x.$$

□

If any of these equivalent conditions hold then the stationary policy is regular.

For a well designed policy the controlled chain will be stable in a far stronger sense. A Markov chain $\boldsymbol{Q}$ is called V*-uniform ergodic*, with $V\colon\mathbb{R}^\ell\to[1,\infty)$ a given function, if there exists $\gamma<1$, and $b<\infty$ such that

$$|\mathsf{E}[g(Q(k))\mid Q(0)=x]-\pi(g)| \le b\gamma^k V(x), \qquad k\in\mathbb{Z}_+,\ x\in\mathsf{X},$$

where g is any function satisfying $|g(x)|\le V(x)$, $x\in\mathsf{X}$ (see [47, Chapter 17]). A Markov chain satisfying this strong form of ergodicity is similar to an i.i.d. processes. In particular, a V-uniform Markov chain satisfies a strong form of the large deviations principle [3, 33].

This stronger form of stability holds under uniform convergence to the fluid limit. The following two forms of uniform convergence will be assumed on a given set $S\subset\mathsf{X}$ of initial conditions. For a set $Y\subseteq\mathbb{R}^\ell$ and a point $x\in\mathbb{R}^\ell$ we define

$$d\{x,Y\}=\inf(\|x-y\|:y\in Y)$$

Similarly, if $\mathcal{F}\subseteq C([0,T],\mathbb{R}^{\ell+\ell_u})$ is a set of functions, and $q\in C([0,T],\mathbb{R}^{\ell+\ell_u})$ is another function, then we define

$$d\{q,\mathcal{F}\}=\inf_{\psi\in\mathcal{F}}\sup_{0\le t\le T}\|q(t)-\psi(t)\|.$$

(U1) For any given T,ε there is a sequence $\{\Theta(\varepsilon,T,n)\}$ such that for any $x\in S$,

$$\mathsf{P}\Bigl(d\{q^n(T;x),Y_x(T)\}>\varepsilon\Bigr)\le\Theta(\varepsilon,T,n)\to 0, \qquad n\to\infty,$$

where $Y_x(T)=\{q(T;x):q\in\mathcal{L}_x\}\subseteq\mathbb{R}^\ell_+$.

(U2) For any given T,ε there is a sequence $\{\Theta(\varepsilon,T,n)\}$ such that for any $x\in S$,

$$\mathsf{P}\Bigl(d\{q^n(\,\cdot\,;x),\mathcal{L}_x\}>\varepsilon\Bigr)\le\Theta(\varepsilon,T,n)\to 0, \qquad n\to\infty.$$

These conditions are frequently automatic since the functions $\{q^n(\,\cdot\,;x) - q(\,\cdot\,;x), z^n(\,\cdot\,,x) : n \geq 1, x \neq \theta\}$ are uniformly bounded, and uniformly Lipschitz continuous.

A stable fluid limit model always admits a *Lyapunov function* $V : \mathbb{R}_+^\ell \to \mathbb{R}_+$ satisfying

$$V(q(t;x)) \leq V(x) - t \qquad \text{for } t < \tau_\theta = \min(t : q(t;x) = \theta). \tag{2.15}$$

One can take the maximal emptying time itself and, moreover, this Lyapunov function is radially homogeneous. Conversely, the existence of a Lyapunov function satisfying (2.15) is known to imply stability.

If there exists a *Lipschitz continuous* Lyapunov function then one can deduce not just stability, but robustness with respect to parametric perturbations. It is not surprising then that the existence of a Lipschitz Lyapunov function implies of form of exponential stability.

Theorem 4. *Suppose that the network is controlled using a stationary policy, and that there exists* $b_0 < \infty$ *such that with* $S = \{x : \|x\| \geq b_0\}$,

(i) *Assumption (U1) holds on* S.

(ii) *There exists a Lipschitz continuous function* $V : \mathbb{R}_+^\ell \to \mathbb{R}_+$ *satisfying (2.15) for* $x \in S$.

Then the network is V_ε*-uniformly ergodic, where* $V_\varepsilon(x) = \exp(\varepsilon V(x))$ *for some* $\varepsilon > 0$ *sufficiently small.*

Proof. Note first of all that V can be taken radially homogeneous without loss of generality: $V(bx) = bV(x)$ for $b \geq 0$. If this is not the case, we can replace V by $V^1(x) = \inf_{b>0} \frac{1}{b} V(bx)$.

Given the uniform convergence of $\{q^n\}$, and the inequality (2.15) for the limit, we can find an $n_0 > 0$ such that

$$\mathsf{E}[V(Q(n;nx))] = n\mathsf{E}[V(q^n(1;x))] \leq n(x - 1/2), \qquad \|x\| \geq b_0,\ n \geq n_0.$$

Hence we can assume that (V1) of [47] is satisfied for the n-step chain:

$$P^n V\,(x) \leq V(x) - 1, \qquad \|x\| \geq nb_0.$$

The result then follows from [47, Theorem 16.3.1]. □

A strengthened form of convergence to the fluid limit model also provides a basis for establishing transience of a network model. Note that a large deviations bound would provide a rate of convergence far stronger than assumed in (i).

Theorem 5. *Suppose that the network is controlled using a stationary policy, and that the following hold for the set* $S = \mathcal{O}$, *where* $\mathcal{O}$ *is bounded, and open as a subset of* $\mathbb{R}_+^\ell$.

(i) *The uniform limit (U2) holds, where for some finite* $b(\,\cdot\,)$,

$$\Theta(\varepsilon, T, n) \leq b(\varepsilon, T)/n, \qquad n \geq 1.$$

(ii) *There is an open set* $\mathcal{V}$ *with* $\bar{\mathcal{V}} \subset \mathcal{O}$, *and an* $r > 1$, $T < \infty$, *such that*

$$q(T;x) \in r\mathcal{V}, \qquad x \in \mathcal{O},\ q \in \mathcal{L}_x.$$

(iii) *For some* $\varepsilon_1 > 0$ *we have the uniform lower bound,*

$$\|q(t;x)\| \geq \varepsilon_1, \qquad 0 \leq t \leq T, x \in \mathcal{O},\ q \in \mathcal{L}_x.$$

Then there is a constant b_2 *such that for* $x \in \mathcal{O}$,

$$\mathsf{P}(\boldsymbol{Q} \to \infty \mid Q(0) = nx) \geq 1 - \frac{b_2}{n}.$$

Hence if $\{n\mathcal{O}\} \cap \mathcal{C} \neq \emptyset$ *for some* $n > b_2$ *then the state process* $\boldsymbol{Q}$ *is a transient Markov chain.*

Proof. By (U2) we have for some $0 < \varepsilon_2 < \varepsilon_1$, some $b_1 < \infty$, and any $x \in \mathcal{O}$,

$$\begin{aligned}
&\mathsf{P}\Big(\|Q(nt;nx)\| \geq \varepsilon_2 n,\ 0 \leq t \leq T,\ \text{and } Q(nT;nx) \in nr\mathcal{O}\Big) \\
&= \mathsf{P}\Big(\|q^n(t;x)\| \geq \varepsilon_2,\ 0 \leq t \leq T,\ \text{and } q^n(T;x) \in r\mathcal{O}\Big) \\
&\geq 1 - \frac{b_1}{n}.
\end{aligned}$$

Here we are also using the assumption that $\bar{\mathcal{V}} \subset \mathcal{O}$.

The above bound can be generalized by replacing the integer n with $r^i n$, where $r > 1$ is given in (ii), (again taking integer parts whenever necessary). For any $x \in \mathbb{R}^\ell_+$, and any $i \geq 1, n \geq 1$, define the event $\mathcal{A}(x,i,n) =$

$$\Big\{\|Q(r^i nt;nx)\| \geq \varepsilon_2 r^i n,\ 0 \leq t \leq T,\ \text{and } Q(r^i nT;nx) \in r^{i+1} n\mathcal{O}\Big\}$$

So that by the previous bound, whenever $x \in r^i\mathcal{O}$,

$$\mathsf{P}(\mathcal{A}(x,i,n)) \geq 1 - \frac{b_1}{n} r^{-i}.$$

By stopping the process at the successive times $N_0 = 0$, $N_k = N_{k-1} + nr^{k-1}T$, $k \geq 1$, and using the Markov property we find that for $x \in \mathcal{O}$, and with $\beta = \varepsilon_2 T^{-1}(1 - r^{-1})$,

$$\begin{aligned}
\mathsf{P}(\|Q(k;nx)\| \geq \beta t,\ 0 \leq t \leq nN_k) &\geq \sum_{i=0}^{k-1} \inf_{y \in \{r^i\mathcal{O}\}} \mathsf{P}(\mathcal{A}(y,i,n)) \\
&\geq 1 - \frac{b_1}{n}(1 + r^{-1} + \cdots + r^{-k+1}).
\end{aligned}$$

This proves the theorem with $b_2 = \frac{b_1}{1-r^{-1}}$ since the integer k is arbitrary. □

We consider the example illustrated in Figure 10 to show how Theorems 4 and 5 can be applied. This example was introduced in [52, 35] to show how instability can arise in networks even when the traffic conditions are satisfied.

To give one example of a stabilizing policy, suppose that at each time k we choose $U^\circ(k)$ to minimize the conditional mean,

$$U^\circ(k) = \arg\min_a \mathsf{E}[\|Q(k+1)\|^2 \mid Q(k), U(k) = a]$$

where the minimum is over all $a \in \mathsf{U}$, subject to the constraint that $a_i = 0$ if $Q_i(k) = 0$. The minimization can be selected so that $\mathbf{U}^\circ$ is defined by a non-idling, stationary Markov policy defined by a feedback law F°.

The feedback law can be written as

$$F^\circ(x) = \arg\min P_a V_2\,(x), \qquad x \in \mathsf{X},$$

where $V_2(\,\cdot\,) := \|\,\cdot\,\|^2$. The function $F^\circ\colon \mathsf{X} \to \mathsf{U}$ is radially constant, and vanishes on the boundaries, $F_i^\circ(x) = 0$ when $x_i = 0$. The uniform condition (U2) is readily verified in this case since, for large x, the controlled chain resembles an unreflected random walk (see [6], and also Proposition 8 below).

To evaluate F° note that for any action a, the drift $P_a V_2 - V_2$ is the sum of a linear term $\langle v_a, x\rangle$ and a bounded term. The vector v_a can be expressed,

$$v_a = 2(Ba + \alpha) = 2(\alpha_1 - \mu_1 a_1, \mu_1 a_1 - \mu_2 a_2, \alpha_3 - \mu_3 a_3, \mu_3 a_3 - \mu_4 a_4)^{T}$$

The choice $a^{\text{ss}} = (\alpha_1\mu_1^{-1}, \alpha_1\mu_2^{-1}, \alpha_3\mu_3^{-1}, \alpha_3\mu_4^{-1})^{T}$ makes $v_{a^{\text{ss}}} = 0$, and is in the interior of the control space provided that the capacity conditions hold. This gives $P_{a^{\text{ss}}} V_2 - V_2 \le b_0$ for some constant b_0. This is a randomized action, which is feasible provided $x_i \neq 0$, $1 \le i \le 4$. If some $x_i = 0$ then the corresponding value a_i^{ss} must also be set to 0, but we still obtain an upper bound of the form $P_{a^{\text{ss}}} V_2 - V_2 \le b_0$.

One can conclude that the feedback law

$$F^\varepsilon(x) = \mathbb{I}_+(x)(a^{\text{ss}} - \varepsilon B^{-1}\frac{x}{\|x\|}) \tag{2.16}$$

with B given in (2.12) and $\mathbb{I}_+(x) = \text{diag}(\mathbb{I}(x_1 > 0), \dots, \mathbb{I}(x_4 > 0))$ is feasible for $\varepsilon > 0$ sufficiently small. For some possibly larger b_0, it satisfies

$$P_{F^\varepsilon} V_2 \le V_2 - 2\varepsilon\|x\| + b_0.$$

By minimality, the feedback law F° exhibits an even larger negative drift,

$$P_{F^\circ} V_2 \le P_{F^\varepsilon} V_2 \le V_2 - 2\varepsilon\|x\| + b_0.$$

Using Jensen's inequality we find that the function $V(x) = \sqrt{V_2(x)} = \|x\|$ is a Lyapunov function for the network, and it is also a Lyapunov function for the fluid limit model. Applying Theorem 4 we see that F° is a regular policy, and that the controlled network is V_ε-uniformly ergodic.

Suppose that replace the ℓ_2-norm by the ℓ_1 norm. Letting $c(x) = \sum x_i$, we minimize over all a the conditional mean $P_a c\,(x)$. The resulting policy is the last buffer-first served policy (where buffers 2 and 4 have strict priority). This is also one version of the $c\mu$-rule.

This policy is known to lead to a *transient* model for certain parameters, even when (2.13) holds. Specifically, suppose that

$$\frac{\alpha_1}{\mu_2} + \frac{\alpha_2}{\mu_4} > 1.$$

Under the 'LBFS' policy the resulting fluid limit model satisfies for some T, r,

$$q(T; x_0) = r x_0 \qquad x_0 = (0,0,0,1)^{T}.$$

This was first shown in [35]. We also have $\|q(t; x_\varepsilon) - q(t; x_0)\| \le \varepsilon$, $0 \le t \le T$, for all $\varepsilon > 0$ sufficiently small, and all $x_\varepsilon \in \mathbb{R}_+^\ell$ satisfying $\|x_\varepsilon - x_0\| \le \varepsilon$. Hence the assumptions of Theorem 5 are satisfied with $\mathcal{O} = \{x \in \mathbb{R}_+ : \|x - x_0\| < \varepsilon\}$, and $\mathcal{V} = \{x \in \mathbb{R}_+ : \|x - x_0\| < r^{-1}\varepsilon\}$. Condition (U2) holds with $S = \mathcal{O}$, again using the fact that between emptying times at buffers 2 and 4 the process $\boldsymbol{Q}$ is a simple random walk. We conclude that the network model is transient under the LBFS policy.

3. Optimization.

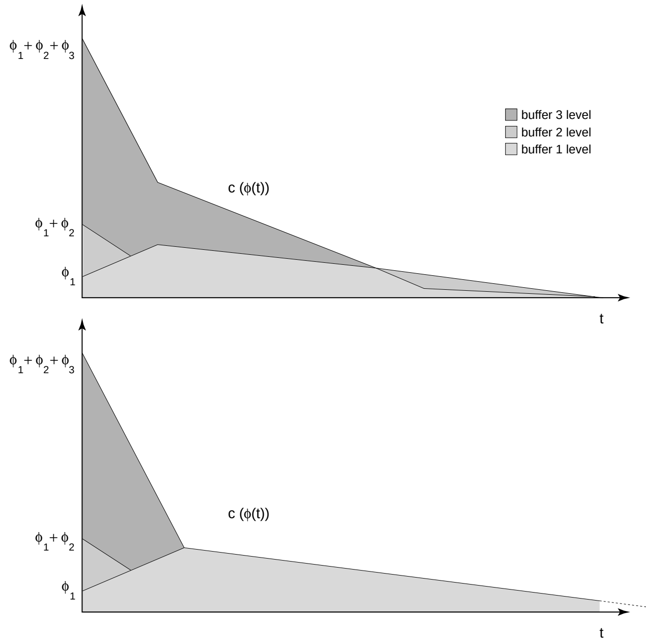

FIG. 3. *The trajectory of buffer levels and evolution of the cost $c(q(t)) = |q(t)|$ for the model of Figure 12 for a given set of initial conditions. The first figure illustrates the optimal policy, and the second shows the last buffer first served priority policy.*

3.1. The average cost optimization problem. In this paper we restrict attention to the average cost problem. For any allocation sequence $\boldsymbol{U}$ and any initial condition x we set

$$J(x, \boldsymbol{U}) = \limsup_{N\to\infty} \frac{1}{N} \sum_{0}^{N-1} \mathsf{E}_x[c(Q(k))].$$

The most common choice is $c(x) = |x|$, where we let $|\cdot|$ denote the ℓ_1 norm. In this case the optimization of J amounts to delay minimization, by Little's Theorem. We have already seen that the cost $J(x, \boldsymbol{U})$ is finite and independent of x when $\boldsymbol{U}$ is a regular policy.

An optimal policy, if it exists, can be taken to be stationary, where the associated average cost optimality equations are given in (1.1,1.2). We show below that a solution does exist when the fluid model is stabilizable, and that h^* can be approximated by the value function for a fluid model optimal control problem. For any T and any $x \in \mathbb{R}_+^\ell$, consider the problem of minimizing

$$\int_0^T c(q(t; x))\, dt$$

subject to the constraint that $q\colon [0, T] \to \mathbb{R}_+^\ell$ satisfy (2.2) for some feasible allocation $\boldsymbol{z}$. The infimum is denoted $V^*(x, T)$.

The following proposition shows that the fluid optimal policy is in some sense *greedy*. That is, the cost as a function of the state $c(q(t))$ is never increasing, and its rate of decrease is maximal when $t \sim 0$. Such behavior is rarely found in dynamic optimization problems. For example, even a second order linear system controlled using optimal (LQR) linear feedback can be expected to exhibit overshoot. An illustration is shown in Figure 3 for the network shown in Figure 12.

The proof of (ii) follows from the aforementioned fact that a state can be reached by following a straight line, provided it is reachable through some control. The result (i) easily follows, and (iii) is well known (see [56, 50]).

PROPOSITION 6. *For any time horizon T,*

(i) *The value function $V^*(\,\cdot\,,T)$ is convex;*

(ii) *For any $x \in \mathbb{R}_+^\ell$, and any optimal allocation $z^*(\,\cdot\,;x)$, the function $c(z^*(t;x))$ is decreasing and convex as a function of t;*

(iii) *If the cost c is linear then $V^*(\,\cdot\,,T)$ is piecewise quadratic. Moreover, for any $x \in \mathbb{R}_+^\ell$ there exists an optimal state trajectory $q^*(\,\cdot\,;x)$ and an optimal allocation $z^*(\,\cdot\,;x)$ which are piecewise linear.*

□

For any fixed x we evidently have that $V^*(x,T)$ is increasing with T. Moreover, there exists some T_θ such that the optimal trajectories vanish by time T_θ when the time horizon T is at least T_θ, whenever the initial condition satisfies $\|x\| \leq 1$. It follows that $V^*(x,T) = V^*(x,T_\theta)$ for any such T and x (see the proof of Theorem 7 (ii) below). Hence for such x and T we have

$$V^*(x,T) = V^*(x) = \min \int_0^\infty c(q(t))\,dt \tag{3.1}$$

where the minimum is subject to the same constraints on q over the entire positive time axis.

THEOREM 7. *If the fluid model is stabilizable, then for the network (2.7) there is a stationary, non-randomized feedback law F^* with the following properties:*

(i) *It is regular, and hence the average cost $\eta^* = J(x,F^*)$ is finite and independent of x.*

(ii) *The fluid limit model $\mathcal{L}^*$ is stable.*

(iii) *The fluid limit model is optimal with respect to the total cost: With probability one, for any $x \in \mathbb{R}_+^\ell$, and any fluid limit $q^* \in \mathcal{L}_x^*$,*

$$\int_0^\infty c(q^*(t;x))\,dt = V^*(x).$$

(iv) *There exists a solution h^* to the average cost optimality equation (1.1,1.2) which satisfies*

$$\limsup_{\|x\|\to\infty} \Big| \frac{h^*(x)}{\|x\|^2} - \frac{V^*(x)}{\|x\|^2} \Big| = 0.$$

Proof. Result (i) is a minor generalization of [45, Theorem 5.2]. The existence of a stabilizing policy, as required in this result, is guaranteed by stabilizability of the fluid model (see Theorem 13 below).

To prove (ii), assume that (iii) holds. We then have for all t, with probability one,

$$V^*(q^*(t;x)) = V^*(x) - \int_0^t c(q^*(s;x))\,ds.$$

From the Lipschitz continuity of the model one can show that V^* is equivalent to a quadratic, in the sense that $V^*(x)/\|x\|^2$ is bounded from above and below for $x \neq \theta$

[45]. It then follows that for some $l_0 < \infty$,

$$\sqrt{V^*(q^*(t;x))} \le \sqrt{V^*(x)} - l_0 t, \qquad t \le \tau_\theta,$$

where τ_θ is the emptying time for $q^*(t;x)$. Thus we have the bound $\tau_\theta \le \sqrt{V^*(x)}/l_0 < \infty$

To prove (iii) we use the following previous results:

(a) It is shown in [45, Theorem 5.2 (i)] that the policy F^* can be chosen so that for any other policy F,

$$\liminf_{T\to\infty} \liminf_{n\to\infty} \Big(\mathsf{E}\Big[\int_0^T c(q^n_F(t;x))\, ds\Big] - \mathsf{E}\Big[\int_0^T c(q^n_{F^*}(s;x))\, ds\Big]\Big) \ge 0. \tag{3.2}$$

(b) In [42, 2] it is shown that for any T there is a policy F^∞ which attains the optimal cost for the fluid control problem:

$$\lim_{n\to\infty} \mathsf{E}\Big[\int_0^T c(q^n_{F^\infty}(s;x))\, ds\Big] = V^*(x,T), \qquad \|x\| = 1.$$

Taking $T > T_\theta$ we see that, for any $m \ge 1$,

$$\lim_{n\to\infty} \mathsf{E}\Big[\int_0^{mT} c(q^n_{F^\infty}(s;x))\, ds\Big] = V^*(x), \qquad \|x\| = m.$$

The point of (b) is that the optimal cost $V^*(x)$ is attainable, and hence the bound given in (a) can be strengthened: For any weak limit $q_{F^*}(\,\cdot\,;x)$ and any T we have by weak convergence,

$$\mathsf{E}\Big[\int_0^T c(q_{F^*}(s;x))\, ds\Big] \le V^*(x), \qquad x \in \mathbb{R}^\ell_+.$$

But for $T \ge T_\theta\|x\|$ we have, with probability one,

$$\int_0^T c(q_{F^*}(s;x))\, ds \ge V^*(x).$$

Combining these two inequalities completes the proof of (iii).

Result (iv) then follows as in the proof of [45, Theorem 5.2 (iii)]. □

From these results we can obtain much insight into the structure of optimal policies for the discrete network when the state is large, i.e., the network is congested. We illustrate this now with several examples.

3.2. Examples. At this stage in the theory there are no general results which are as striking as can be found in numerical examples. The general principle appears to be that an optimal policy for the discrete network is equal to the fluid policy, suitably modified along the boundary of the state space. In fact, in several special cases it has been shown that an approximately optimal policy can be constructed in this manner [32, 26, 4], and a general approach is developed in [43].

In computing optimal policies for the examples below we are forced to truncate the state space to obtain a finite Markov decision process model. Optimization is still difficult due to the large state spaces involved. For example, for a network with four buffers of size twenty each, the state space contains $m = 160{,}000$ elements, and computing the optimal policy involves inverting an $m \times m$ matrix.

However, one can use successive approximation to obtain a sequence of approximations $\{h_n : n \geq 0\}$. This is also known as the value iteration algorithm. Theorem 7 (iv) suggests an initialization for the algorithm: $h_0 = V^* \approx h^*$. Numerical results obtained in [9] show that this choice can speed convergence by orders of magnitude. We have used this approach in all of the examples below.

Boundary effects for a truncated model can be severe. For instance, for a loss model, if a buffer is full then it may be desirable to serve an upstream buffer: the resulting overflow will reduce the steady state cost. The policies are shown in a truncated region since this behavior has nothing to do with real network dynamics. For example, Figure 5 below shows the optimal policy $F^*(x)$ for all x satisfying $\|x\|_\infty < 25$. In this example the value iteration algorithm was used with a network model allowing 39 customers at each buffer. The dimension of the resulting average cost optimality equations (1.1,1.2) was 40^3 since there are three buffers in this example.

The M/M/1 queue. Recall that the fluid limit model satisfies $q(t;x) = x - \mu z(t) + \alpha t$, $t \geq 0$, where $\alpha + \mu = 1$. Using the notation defined in Section 2 we have

$$B = -\mu, \ C = 1, \quad \text{and} \quad \mathbf{1} = 1.$$

The non-idling policy is given by $\zeta(t) = \frac{d}{dt} z(t) = 1$ when $q(t;x) > 0$. It is optimal for any monotone cost function.

For the discrete-stochastic model with cost $c(x) = x$, the relative value function h^* is given by

$$h^*(x) = \tfrac{1}{2} \frac{x^2 + x}{\mu - \alpha}.$$

The fluid value function is given by,

$$\begin{aligned} V^*(x) \quad &:= \quad \int_0^\infty q(t;x)\, dt \\ &= \quad \tfrac{1}{2} \frac{x^2}{\mu - \alpha}. \end{aligned} \tag{3.3}$$

We see that the error in the approximation $h^*(x) \approx V^*(x)$ is linear in this special case.

The 'criss-cross' network. Figure 4 shows a model introduced in [27] to illustrate the use of Brownian motion approximation for networks. The system parameters are

$$B = \begin{bmatrix} -\mu_1 & 0 & 0 \\ \mu_1 & -\mu_2 & 0 \\ 0 & 0 & -\mu_3 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

and in this model we take $\alpha_2 = 0$. This model has become a standard example.

Optimal policies for the fluid model are easily computed. Take c equal to the ℓ_1 norm, and suppose that $\mu_2 < \mu_3 < \mu_1$. In this case strict priority is given to buffer 3 whenever buffer 2 is non-empty. When this buffer does empty, then the optimal policy sets

$$z_1(t) = \mu_2/\mu_1 \qquad z_3(t) = 1 - z_1(t),$$

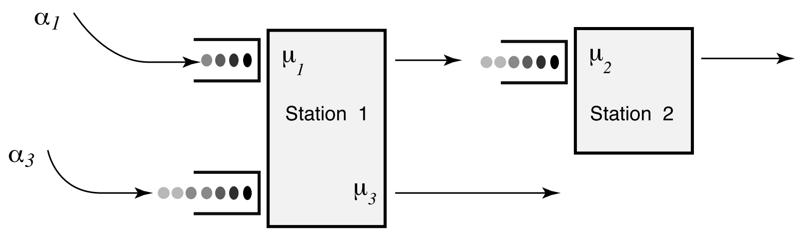

FIG. 4. *A simple two station network with* $\ell_m = 2$ *and* $\ell = \ell_u = 3$.

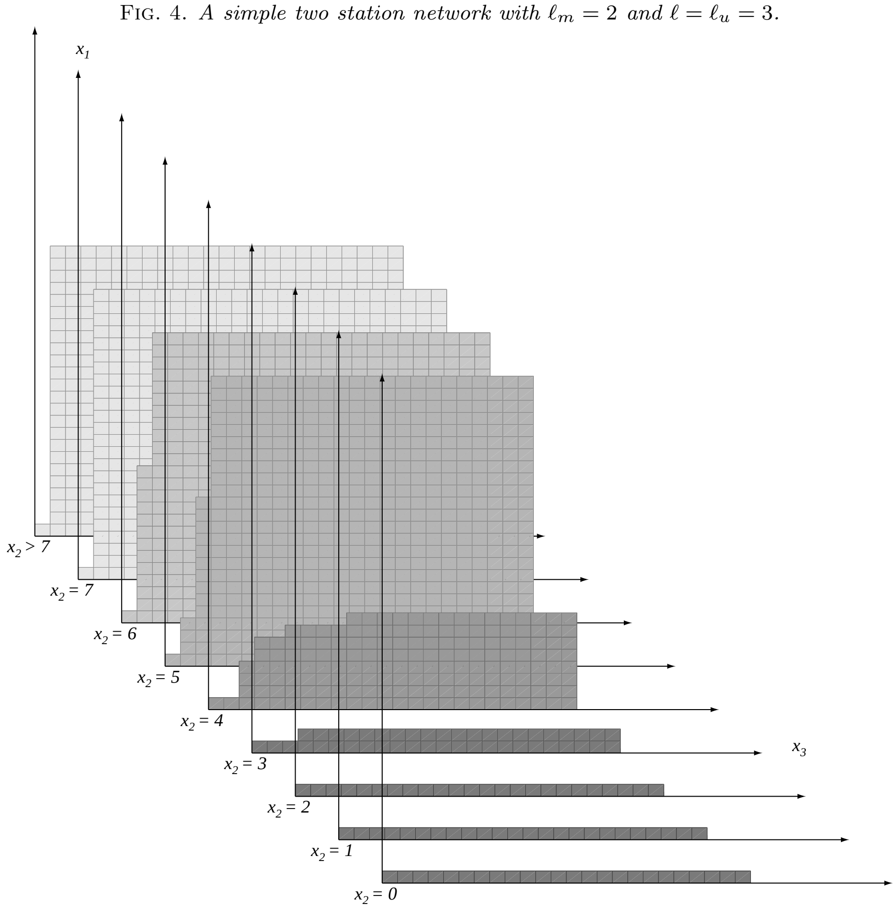

FIG. 5. *The optimal policy for the network above with* $\alpha^T = (9, 0, 9)$ *and* $\mu^T = (25, 10, 20)$. *The grey areas indicate states at which buffer three is given strict priority.*

provided both buffer 1 and buffer 3 are non-empty. This is a path-wise optimal policy in the sense that it minimizes $c(q(t;x))$, *for each* $t \geq 0$, over all policies.

The optimal policy for the discrete model with particular parameter values satisfying these constraints is given in Figure 5. As always, the fluid limit of this optimal policy is the optimal policy for the fluid model. The discrete optimal policy is simlar to the optimal fluid policy: The critical value $q_2(t) = 0$ has been shifted upwards to $Q_2(t) \approx 4$.

A routing model. We now show how the theory applies to a routing problem. The model illustrated in Figure 6 has been considered in several papers, see in particular [23, 32]. Customers that arrive to the system are routed to one of the two servers. In this example $\ell = 2$, $\ell_u = 4$, and

$$B = \begin{bmatrix} \alpha & 0 & -\mu_1 & 0 \\ 0 & \alpha & 0 & -\mu_2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The router is non-idling in this model, $\zeta_1 + \zeta_2 = 1$. This requirement can be expressed as the additional linear constraint, $C_a \zeta \le b_a$, where $C_a = [-1, -1, 0, 0]$, and $b_a = -1$. Alternatively, one can enlarge the state space to include a buffer at the router, but impose the linear constraint that $Q_3(t) \equiv 0$.

Note that in the routing model the arrival stream is absorbed into the random matrix $\widetilde{B}$. Hence, in this model we take the two dimensional vector α to be zero. Assume that $\mu_1 = \mu_2 = \mu$, and that $\mu < \alpha < 2\mu$. The fluid model is then stabilizable.

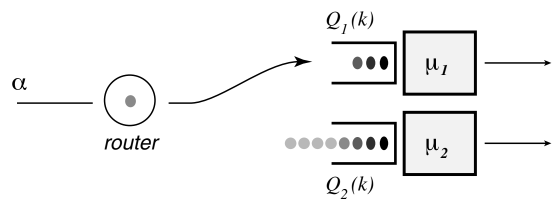

FIG. 6. *A network with controlled routing:* $\ell_m = 3$, $\ell = 2$, *and* $\ell_u = 4$.

To minimize the total cost for the fluid model

$$\int_0^\infty c(q(t;x))\,dt,$$

one obviously takes z_3 and z_4 to be non-idling. Consider the case where c is linear, with $c(x) = (c_1, c_2) \cdot x$, and $c_1 > c_2$. Then the priority policy is optimal, where fluid is routed to buffer two as long as buffer one is non-empty. As soon as it does empty, then fluid is routed to buffer one at rate $\zeta_1(t) = \frac{d}{dt} z_1(t) = \mu_1/\alpha$ so that buffer one is non-idling, but empty. The remaining fluid is sent to buffer two so that $\zeta_2(t) = 1 - \mu_1/\alpha$. This policy is again path-wise optimal, and it enforces non-idleness so that $\zeta_3(t) = \zeta_4(t) = 1$ for $t < \tau_\theta$.

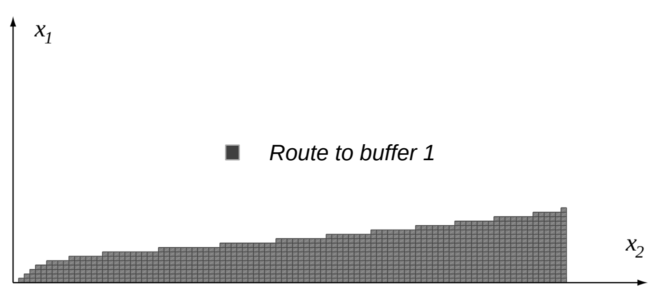

FIG. 7. *Optimal discrete policy for the simple routing model with* $\alpha = 9$ *and* $\mu^T = (5, 5)$. *The one step cost is* $c(x) = 3x_1 + 2x_2$.

The discrete-stochastic model is considered in [23] for a general linear cost function. It is shown that an optimal policy exists, and that it is of a nonlinear threshold form: There is a non-decreasing function $\gamma\colon \mathbb{Z}_+ \to \mathbb{Z}_+$ such that when a job arrives when the queue lengths are x_1, and x_2, then buffer one receives the job if and only if $\gamma(x_1) \geq x_2$. The analysis of [58] implies that the function γ is unbounded, but in general no analytic formula is available for the computation of γ.

We see in Figure 7 that the optimal policy for the discrete network is closely approximated by the optimal fluid policy, modified along the boundary. The 'thickened boundary' ensures that, with high probability, neither buffer will idle when the network is congested.

A processor-sharing model. Another simple example where the optimal allocation for the fluid model is explicitly computable is the processor sharing network considered in [4, 26], illustrated in Figure 8. In this example $\ell = 2$, $\ell_u = 3$, and the system parameters are

$$B = \begin{bmatrix} -\mu_A & -\mu_B & 0 \\ 0 & 0 & -\mu_C \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_A \\ \alpha_B \end{bmatrix}$$

As in the previous example, for any cost c we can assume that the optimal policy is non-idling at station 1. The fluid limit model illustrated on the right in Figure 8 is based on this assumption.

We assume that $\alpha_A > \mu_A$. In this case it is critical that Station 1 receive outside assistance. Under this condition and the non-idling assumption at Station 1, we arrive at a reduced order model with $\ell = \ell_u = 2$, and

$$B = \begin{bmatrix} -\mu_B & 0 \\ 0 & -\mu_C \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_A - \mu_A \\ \alpha_B \end{bmatrix}$$

For any linear cost the optimal allocation is the $c\mu$-rule priority policy, which is again path-wise optimal in this example. It is shown in [4] that a modification of this policy is nearly optimal in heavy traffic. Figure 9 shows the optimal policy for the discrete model. It is similar to the $c\mu$ priority policy, with priority given to processor B at station 2. However the boundary $\{x_2 = 0\}$ has been shifted to form the concave region shown in the figure.

A generalized $c\mu$-rule in scheduling. We return now to the example illustrated in Figure 10, whose fluid model is defined by the parameters

$$B = \begin{bmatrix} -\mu_1 & 0 & 0 & 0 \\ \mu_1 & -\mu_2 & 0 & 0 \\ 0 & 0 & -\mu_3 & 0 \\ 0 & 0 & \mu_3 & -\mu_4 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

with $\alpha_2 = \alpha_4 = 0$. Consider for simplicity the symmetric case where $\mu_1 = \mu_3$, and $\mu_2 = \mu_4$. We also assume that $\mu_1 = 2\mu_2$ so that the exit buffers are slow.

It is pointed out in [42] that the optimal policy for the fluid model when c is the ℓ_1 norm is given as follows: The exit buffers have strict priority when $q_2(t) > 0$ and $q_4(t) > 0$. As soon as one of these buffers empties, say buffer two, then one sets $\zeta_1(t) = \mu_2/\mu_1$, $\zeta_4(t) = 1 - \mu_1$, and continue to set $\zeta_2(t) = 1$. This policy maximizes the overall draining rate at each time t, it is path-wise optimal, and it achieves the total cost V^* for the fluid model. Recall that the analogous greedy policy, defined through the discrete model, is destabilizing!

We have computed an optimal policy for the discrete network numerically for this special case. However, noting that the optimal fluid policy is independent of

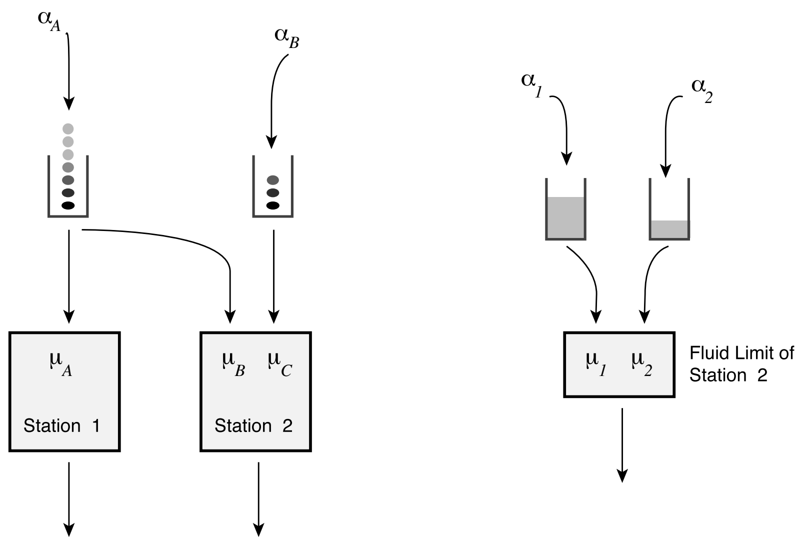

FIG. 8. *On the left is the processor-sharing network of [27]. On the right is its fluid limit model.*

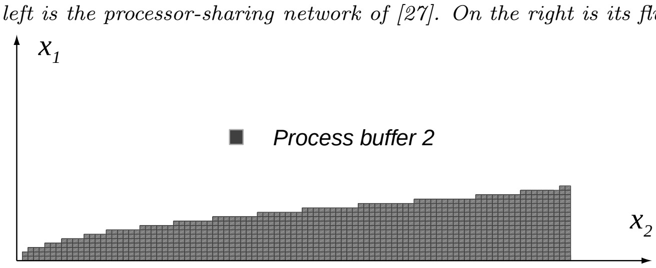

FIG. 9. *Optimal policy for the processor sharing model with ℓ_1 cost, $\alpha = (1, 1)$, and $\mu = (1, 3, 2)$. This closely resembles the optimal fluid policy which gives strict priority to the first server at station two since $\mu_B > \mu_C$.*

the arrival rates $\alpha^T = (\alpha_1, 0, \alpha_3, 0)$, we have taken an extreme case with $\alpha = 0$, and consider the total cost problem,

$$V(x) = \min \sum_0^\infty \mathsf{E}_x[c(Q(k))],$$

where the minimum is with respect to all policies. This gives rise to a finite dimensional optimization problem which can be solved exactly for each x.

We again see that the optimal discrete policy is similar to the optimal fluid policy.

A scheduling-model with no path-wise optimal solution. Consider the network given in Figure 12 with c taken to be the ℓ_1 norm. One policy that minimizes the total cost for the fluid model is defined as follows, where γ is a positive constant defined by the parameters of the network.

(i) Serve $q_3(t)$ exclusively ($\zeta_3(t) = 1$) whenever $q_2(t) > 0$ and $q_3(t) > 0$;
(ii) Serve $q_3(t)$ exclusively whenever $q_2(t) = 0$, and $q_3(t)/q_1(t) > \gamma$;
(iii) Give $q_1(t)$ partial service with $\zeta_1(t) = \mu_2/\mu_1$ whenever $q_2(t) = 0$, and

$$0 < q_3(t)/q_1(t) \le \gamma$$

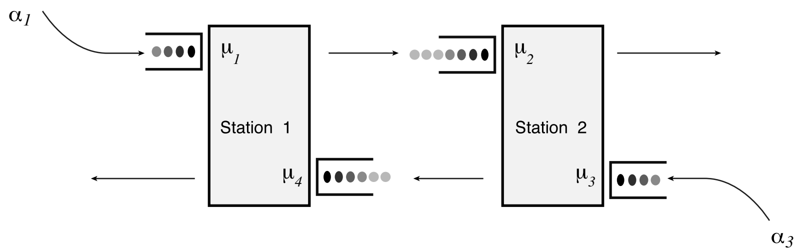

FIG. 10. *A multiclass network with* $\ell_m = 2$ *and* $\ell = \ell_u = 4$.

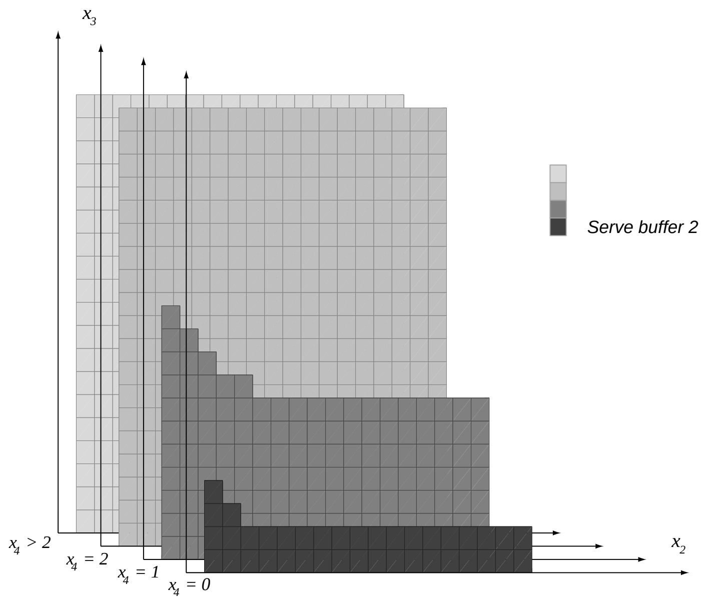

FIG. 11. *Optimal policy for the four-buffer scheduling model shown in Figure 10 under the total cost criterion, with* c *equal to the* ℓ_1 *norm. The arrival streams are null, and* $\mu = (2, 1, 2, 1)$*. The figure shows the policy when* $x_1 = 3$*,* $x_4 = 0, 1, 2, 3$*, with* x_2 *and* x_3 *arbitrary. This optimal policy is of the same form as the fluid policy: It gives strict priority to the exit buffer at station 2, unless buffer 4 is starved of work, in which case buffer 3 releases parts to feed buffer 4.*

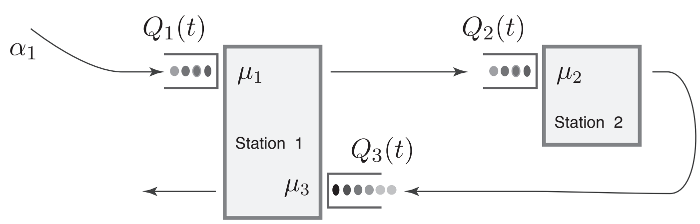

FIG. 12. *A multiclass network with* $\ell_m = 2$ *and* $\ell = \ell_u = 3$.

This model is most interesting when station two is the bottleneck, since one must then make a tradeoff between draining the system, and avoiding starvation at the bottleneck. Taking $\rho_2 = \alpha/\mu_2 = 9/10$, and $\rho_1 = \alpha/\mu_1 + \alpha/\mu_3 = 9/11$, the constant γ is equal to one, and hence the optimal policy is of the form illustrated in Figure 13.

Optimal policies are computed numerically in [45] for versions of this model with truncated buffers. Results from one experiment are shown in Figure 14. As in the previous examples we see that the discrete optimal policy is easily interpreted. It regulates the work waiting at buffer 2, and does so in such a way that buffer 2 is rarely starved of work when the network is congested.

The policy shown in Figure 14 is also very similar to the fluid policy. Performing some curve fitting, we can approximate this discrete policy as follows: serve buffer one at time t if and only if either buffer three is equal to zero, or

$$Q_1(k) - \bar{x}_1 > Q_3(k) - \bar{x}_3 \qquad \text{and } Q_2(k) \leq \bar{x}_2, \tag{3.4}$$

where the translation $\bar{x}$ positive. The most accurate approximation is obtained when $\bar{x}$ depends upon the current state $Q(k) = x$, say

$$\bar{x} = \bar{x}_0 \log(\|x\| + 1), \qquad x \in \mathsf{X}, \tag{3.5}$$

with $\bar{x}_0 > 0$ and constant. Moreover, with this choice, the fluid limit obtained using the policy (3.4) is precisely the optimal policy minimizing the total fluid cost which is illustrated in Figure 13.

4. Feedback regulation. The results and examples of the previous section all suggest that the fluid model should play a useful role in control synthesis for network models. Theorem 7 establishes a connection between two optimization problems: One is deterministic and relatively transparent; the other stochastic, discrete, and apparently hopeless.

Even if one can find a feedback law $\frac{d}{dt}z(t) = \zeta(t) = f^*(q(t))$ which is optimal for the fluid model, it is not obvious how to use this information. A direct translation such as $F(x) = f^*(x)$ is not appropriate. It is shown in [45] that this policy may have a fluid limit model which differs grossly from the desirable optimal fluid process. However the numerical results given above all show that, at least for simple models, an optimal policy for a discrete network is approximated by an affine shift of the form

$$F(x) = f^*((x - \bar{x})^+), \qquad x \in \mathsf{X}. \tag{4.1}$$

Moreover, one can show that a properly defined shift of this form ensures that the resulting fluid limit model for the network controlled using F approximates the optimized fluid model (see [42, 2, 43] and Proposition 8 below).

One might arrive at the policy (4.1) without any consideration of optimization. When the feedback law f^* is chosen appropriately, this policy will attempt to regulate the state $Q(k)$ about the value $\bar{x}$. If this regulation is accomplished successfully then,

(i) Provided $\bar{x}$ is not too big, the cost $c(Q(k))$ will not be too large;

(ii) If the target $\bar{x}$ is not too small then this policy will avoid starvation of any resource; and

(iii) Regulation to a constant should provide reduced variance at each station, as has been argued for the class of fluctuation smoothing policies [37].

4.1. Discrete review structure. In this section we adapt the approach of [28] to define policies for the physical network based on an idealized allocation derived from the fluid model. Related approaches are described in [22].

In practice one will rarely use a stationary policy F since one is forced to make scheduling decisions at every discrete sampling instance. This is undesirable since it results in high computational overhead, and more importantly excessive switch-overs. The proposed policies consist of three components:

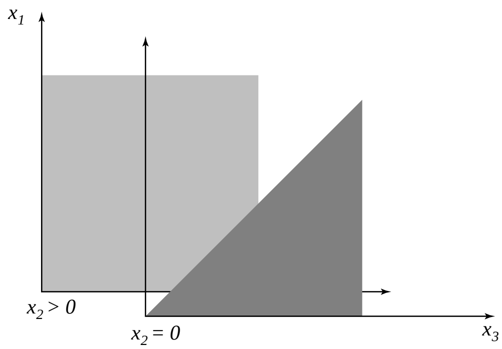

FIG. 13. *The optimal fluid policy for the three buffer re-entrant line with* $\rho_2 = 9/10$ *and* $\rho_1 = 9/11$. *In this illustration, the grey regions indicate those states for which buffer three is given exclusive service.*

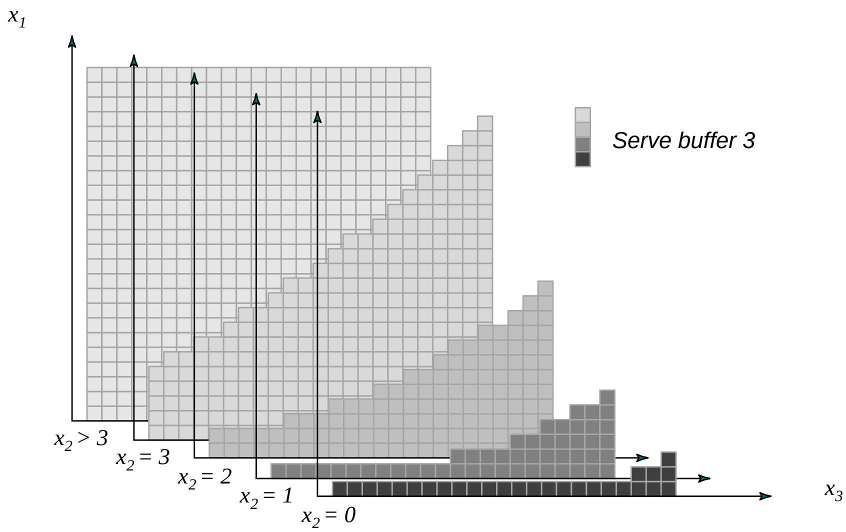

FIG. 14. *Optimal discrete policy for three buffer re-entrant line in the balanced case:* $\alpha/\mu_1 = \alpha/\mu_3 = \frac{1}{2}\rho_1$.

(a) For each initial condition x, a well designed fluid trajectory $q(t;x)$ satisfying (2.2) for some allocation process $\boldsymbol{U}$. This will typically be defined through a feedback law f so that

$$q(t;x) = x + \alpha t + B\int_0^t f(q(s;x))\,ds, \qquad t \in \mathbb{R}_+,\ x \in \mathbb{R}_+^\ell.$$

(b) A target vector $\bar{x}$;

(c) A time horizon N over which the policy is fixed.

The target $\bar{x}$ may be a 'moving target', in which case it is assumed to be a function of the state. In general we may also take N as a function of the state, but we will always assume that N and $\bar{x}$ are 'roughly constant' for large x.

Given these, we set $N_0 = 0$ and, given $Q(N_0) = x$, we determine the time horizon $N_1 = N(x)$. The values $U(k)$, $N_0 \le t < N_1$, are chosen so that

$$\mathsf{E}[Q(N_1 - N_0; x)] - x \approx \delta(N_1 - N_0; (x - \bar{x})^+) \tag{4.2}$$

where $\delta(T;y) = q(T;y) - y$ for any $y \in \mathbb{R}_+^\ell$, $T \geq 0$.

This final choice is far from unique, but will be dictated by considerations such as minimizing switch-over times and avoiding starvation at any station. Once one arrives at time N_1, the choice of $U(k)$ on the time interval $[N_1, N_2)$ proceeds exactly as before, where $N_2 = N_1 + N(Q(N_1))$. Successive review times $\{N_i : i \geq 0\}$ and actions $\{U(k) : k \geq 0\}$ can then be found by induction.

We shall call any policy of this form a *feedback regulation policy* since it is similar to a state feedback approach to regulation as covered in a second year control systems course. Below we list some of the issues in design:

- The most basic question is the design of the fluid state trajectories $\{q(\,\cdot\,;x) : x \in \mathbb{R}_+^\ell\}$. A first requirement is stability, and the theory suggests that good performance for the fluid model with respect to the total cost is highly desirable. These design issues will be discussed in depth in Section 4.2.
- How do we choose $\bar{x}$?
- The time horizon N? This will be dictated by such issues as batch or packet sizes, and switch-over costs.
- How do we choose a sequence of actions on $[N_k, N_{k+1})$ so that (4.2) holds? Again one must consider switch-over costs - two approaches are described below.
- In many models one must also consider idleness avoidance. For routing models this can be treated as in [32] by enforcing routing to a buffer whenever its level falls below a threshold. Scheduling models can be treated similarly.
- Machine failures and maintenance: How should the policy change during a failure? One can again address this problem by considering a fluid model, but one should consider the *delayed* fluid model which includes the residual-life of each service process.

Much has been written on the choice of safety-stock levels. In some simple examples a constant threshold is optimal (see e.g. [18, 39]). A value of zero is optimal in the single-machine scheduling problem with linear cost since the $c\mu$-rule is optimal for both the fluid and stochastic models.

A general approach is suggested by a rich literature on networks in heavy traffic. In [32] and many other references one considers the case where the system load ρ is close to unity, and

$$(1-\rho)\sqrt{n} \to L, \qquad n \to \infty,$$

where n is a parameter which is sent to ∞ for the purpose of analysis, $\rho = \rho(n)$, and L is a non-zero, finite number. The steady state number of customers in the system is typically of order $(1-\rho)^{-1}$ (consider a $G/G/1$ queue, or the functional bounds obtained in [31]). Hence the assumptions commonly used in the literature imply that

$$\sqrt{n} = O\bigl(\frac{1}{1-\rho}\bigr) = O(\mathsf{E}_\pi[\|Q(k)\|]),$$

where $\mathsf{E}_\pi[\|Q(k)\|]$ denotes the steady state mean. The thresholds $\bar{x}$ determined in [4, 32] are of order $\log(n)$, so that $\|\bar{x}\| = O\bigl(\log(\mathsf{E}_\pi[\|Q(k)\|])\bigr)$. By replacing the steady state mean with the current value $Q(k) = x$ we arrive at $\|\bar{x}(x)\| = O(\log(\|x\|))$.

However, the results of [43] show that such a small offset may be overly-optimistic in general. We are currently exploring parameterizations of the form (3.5) where $\bar{x}_0$ can be tuned, perhaps on-line.

Two major issues are stability and performance of the network under a feedback regulation policy. A stability proof and some approaches to estimating performance through simulation are given in Section 4.4. Stability requires some assumptions:

(A1) There exists a function $V\colon \mathbb{R}_+^\ell \to \mathbb{R}_+$ which is Lipschitz continuous, and satisfies,

$$V(q(t;x)) - V(x) \le -t \qquad x \in \mathbb{R}_+^\ell, t \le \tau_\theta.$$

(A2) There exists an i.i.d. process $\{\Gamma(k) : k \ge 0\}$ evolving $\mathbb{R}^\ell$, and a family of functions $\{\boldsymbol{F}(\,\cdot\,, k, N) : 0 \le k < N < \infty\}$ such that for each $k < N$,

$$\boldsymbol{F}(\,\cdot\,, k, N)\colon \mathsf{X}^{k+1} \times \mathbb{R}^{\ell(k+1)} \to \{0,1\}^{\ell_u}$$

and when $N(Q(0)) = N$,

$$U(k) = \boldsymbol{F}(Q(0), \dots, Q(k), \Gamma(0), \dots, \Gamma(k), k, N).$$

(A3) The convergence to the fluid model is uniform in the sense of (U1): For any fixed N, there is a $\Theta(N) > 0$ and $b_0(N) < \infty$ such that when $U(k) = \boldsymbol{F}(Q(0), \dots, Q(k), \Gamma(0), \dots, \Gamma(k), k, N)$ for $k < N$,

$$\frac{1}{N}\mathsf{E}_x\Big[\|Q(N;x) - q(N,(x-\bar{x})^+)\|\Big] \le \Theta(N), \qquad \|x\| \ge b_0(N),$$

where $\Theta(N) \to 0$ as $N \to \infty$.

Assumption (A1) ensures that the fluid process $q(t;x)$ is stable. The Lipschitz assumption is an important aspect of these policies since it is what allows us to establish robustness with respect to perturbations in system parameters such as arrival and service rates.

The proof of Theorem 13 below is based on a Markovian description of the controlled network, which is possible by Assumption (A2). Under this assumption we can define the Markov state process,

$$\Phi(k)^{\scriptscriptstyle T} = [Q(k), \dots, Q(n(k))], \qquad t \in \mathbb{Z}_+,$$

where $n(k)$ is the last switch-over time: $n(k) = \min(s \le t : s = N_k \text{ for some } k)$.

Assumption (A3) requires that one faithfully follow the fluid model. Consider for simplicity the network scheduling problem where $\ell = \ell_u$. Perhaps the most natural approach is to define a processing plan on $[N(k), N(k+1))$ in a generalized round-robin fashion: Take $k = 0$, without loss of generality, and given $Q(0) = x$ set

$$y = q(T;(x-\bar{x})^+) = (x-\bar{x})^+ + Bz(T;(x-\bar{x})^+) + T\alpha.$$

The vector $a = z(T;(x-\bar{x})^+)/T$ satisfies $a \ge 0$, $Ca \le \mathbf{1}$, and since $y = (x-\bar{x})^+ + T(Ba+\alpha)$ we obviously have,

$$(Ba+\alpha)_i \ge 0 \text{ whenever } x_i \le \bar{x}_i. \tag{4.3}$$

At machine s suppose we have ℓ_s buffers $i_1, \dots, i_{\ell_s}$. Given this value of a, and given a constant $m > 0$, we perform $a_{i_1}m$ consecutive services at buffer i_1; then $a_{i_2}m$ consecutive services at buffer i_2; continuing until buffer i_{ℓ_s} has received $a_{i_{\ell_s}}m$ services; and then returning to buffer i_1. This cycle repeats itself until time $N(1)$. We again must take the integer parts of $a_k m$, and this will lead to some error. This is insignificant for large N.

An approach based on randomization is particularly straightforward. Suppose that the function $\boldsymbol{F}$ is formed in a stationary, randomized fashion as follows:

$$\begin{aligned}&\mathsf{P}(U_i(k) = 1 \mid Q(0), \dots, Q(k), \Gamma(0), \dots, \Gamma(k-1))\\ &= \mathsf{P}(U_i(t) = 1 \mid Q(k))\\ &= a_i \mathbb{I}(Q_i(k) > 0).\end{aligned} \tag{4.4}$$

This construction of $U(k)$ for $t < N$ can be equivalently expressed through a feedback function of the form required in (A2) where, for some fixed function $F\colon \mathsf{X} \times \mathbb{R}^\ell \to \mathbb{R}^\ell_+$,

$$\boldsymbol{F}(Q(0), \dots, Q(k), \Gamma(0), \dots, \Gamma(k), t, N) = F(Q(k), \Gamma(k)).$$

Proposition 8. *Consider the network scheduling problem. The randomized policy given in (4.4) defines a function* $\boldsymbol{F}$ *satisfying (A2) and (A3).*

Proof. We have already seen that (A2) holds.

For any $a \in \mathbb{R}^\ell_+$ we let $\boldsymbol{Q}^a = \{Q(k; x, a) : k \geq 0\}$ denote the state process for a Jackson network with arrival and service rates given respectively by $(\alpha_i, a_i\mu_i)$, $1 \leq i \leq \ell$. We always assume that $a \in \mathsf{U}$ ($a \in \mathbb{R}^\ell_+$ with $Ca \leq \mathbf{1}$). We can construct all of the processes $\{Q(\,\cdot\,; x, a) : x \in \mathsf{X}, a \in \mathsf{U}\}$ on the same probability space as follows: We are given two mutually independent ℓ-dimensional Poisson processes $M(t), N(t)$ of rate one. The length of the service times, either real or virtual, at buffer i are defined as the inter-jump times of $N_i(\mu_i a_i t)$; the exogenous arrivals to buffer i occur at the jump times of $M_i(\alpha_i t)$, $t \in \mathbb{R}_+$. The kth component of $\boldsymbol{U}^a$ is given by $U_k(k; x, a) = \mathbb{I}(Q_k(k; x, a) > 0)$.

For each n, we let $\{q^n(t; x, a), z^n(t; x, a)\}$ denote the nth scaled queue-length and cummulative allocation process. We then set

$$\mathcal{L}_n = \overline{\overline{\bigcup_{k=n}^\infty \Big\{ (q^k(\,\cdot\,; x, a), z^k(\,\cdot\,; x, a)) : \|x\| = b_0; a \in \mathsf{U} \Big\}}}$$

The double bar indicates *strong closure* in the function space $C([0,T], \mathbb{R}^{\ell+\ell_v})$, in the uniform norm. The set $\mathcal{L}_n \subset C([0,T], \mathbb{R}^{\ell+\ell_v})$ is compact for any n, and so is its intersection over all n: $\mathcal{L} = \bigcap_n \mathcal{L}_n$.

The set $\mathcal{L}$ is defined for a.e. sample path of (M, N). If $(q, u) \in \mathcal{L}$ then there exists $x^i \to x$, $a^i \to a$, and a subsequence $\{n^i\}$ of $\mathbb{Z}_+$ such that

$$q^{n^i}(\,\cdot\,; x^i, a^i) \Longrightarrow q, \quad z^{n^i}(\,\cdot\,; x^i, a^i) \Longrightarrow z, \qquad i \to \infty,$$

where the convergence is in $C([0,T], \mathbb{R}^{\ell+\ell_v})$. We then have $q(0) = x$, and for any time t at which q and $\boldsymbol{U}$ are differentiable,

$$\begin{aligned}&\frac{d}{dt} q_i(t) = 0 \qquad \text{if } q_i(t) = 0;\\ &\frac{d}{dt} z_i(t) = 1 \qquad \text{if } q_i(t) > 0, \qquad 1 \leq i \leq \ell.\end{aligned}$$

This is enough to completely determine the limit set: For any (x, a) there is a unique $q(\,\cdot\,; x, a) \in \mathcal{L}$.

It follows that we have uniformity in the sense of (U2): For any $\varepsilon > 0$,

$$\sup_{\substack{\|x\|=b_0 \\ a\in \mathsf{U}}} \mathsf{P}\Big(\sup_{0\le t\le T} \|z^n(t;x,a) - z(t;x,a)\| > \varepsilon\Big) \to 0, \qquad n\to\infty, \tag{4.5}$$

and the analogous limit holds for $\{q^n\}$.

For $x\in\mathbb{R}^\ell_+$, $T>0$, we denote

$$\mathcal{A}(x,T) = \{a\in\mathbb{R}^\ell_+ : Ca \le \mathbf{1}, \quad x + T(Ba+\alpha)\in\mathbb{R}^\ell_+\}.$$

For the randomized policies considered in Proposition 8 we always have $a\in\mathcal{A}(x,T)$ (see the discussion surrounding (4.3)). Note also that for such a we have $z_i(t)=t$ for any i, and any $0\le t\le T$. Hence from (4.5),

$$\sup_{\substack{\|x\|=b_0 \\ a\in \mathcal{A}(x,T)}} \mathsf{P}\Big(\sup_{0\le t\le T} \|z_i^n(t;x,a) - t\| > \varepsilon\Big) \to 0, \qquad n\to\infty.$$

It is well known that Jackson networks are *monotone* in the sense that if $y\ge x$ then $U_i(k;y,a)\ge U_i(k;x,a)$ for any i,t and a (see [53]). Hence the above bound can be improved:

$$\sup_{\substack{\|x\|\ge b_0 \\ a\in \mathcal{A}(x,T)}} \mathsf{P}\Big(\sup_{0\le t\le T} \|z_i^n(t;x,a) - t\| > \varepsilon\Big) \to 0, \qquad n\to\infty, \tag{4.6}$$

and this easily implies that (A3) holds. □

4.2. Design of the fluid trajectory. There are many control strategies for a fluid model which have desirable stability characteristics. Here we describe four approaches which always lead to a stabilizing solution: In each case we construct a Lipschitz continuous Lyapunov function. We have already seen in Theorem 4 that this can imply a strong form of stochastic stability for the network. These results will be generalized to feedback regulation policies in Theorem 13.

The first three classes of policies considered below are based on optimal control: The optimal fluid policies, time-optimal fluid policies, and constrained complexity optimal fluid policies. The latter have fixed complexity which can be chosen in advance by the user. The fourth class that we consider consist of greedy policies. Such policies can be computed easily for large networks by solving an ℓ-dimensional linear program.

Optimal fluid policies. The policy f^* which optimizes the fluid model under the total cost criterion is a natural candidate for application in a feedback regulation policy. The computation of f^* may be posed as an infinite dimensional linear program. Because of the specific structure of the linear program it is frequently feasible to compute f^* numerically, even though such problems are in general intractable (see [41, 49]).

Time-optimal allocations can be computed with only trivial calculation: Proposition 1 implies that a linear time-optimal policy can be constructed for any stabilizable network, which is the basis of the main result of [17]. Time-optimality is used as a *constraint* in the construction of the policies described in [43].

Optimal fluid policies are stabilizing for the fluid model. Moreover they satisfy Assumption (A1) and are hence guaranteed to be stabilizing for the stochastic model when used in a feedback regulation policy.

PROPOSITION 9. *Suppose that the fluid model is stabilizable*

(i) *Suppose that $q(\,\cdot\,;x)$ is optimal with respect to the total cost (3.1) for each x. Then there exists a Lipschitz Lyapunov function so that (A1) holds.*

(ii) *Suppose that $q(\,\cdot\,;x)$ is time-optimal in the sense that $\tau_\theta(x)$ is minimized over all fluid policies. Then $V(x) = \tau_\theta(x)$, $x \in \mathbb{R}^\ell_+$, is a Lipschitz Lyapunov function, so that (A1) holds.*

Proof. Let $V = \beta\sqrt{V^*}$ where $\beta > 0$. From Proposition 6 we can conclude that V is radially homogeneous, and each sublevel set $S_\eta = \{x : V(x) \le \eta\}$ is a convex subset of $\mathbb{R}^\ell_+$ for any $\eta > 0$. It follows that V itself is convex and continuous, which implies Lipschitz continuity. The negative drift required in (A1) holds for β sufficiently large, which establishes (i).

Another Lyapunov function is $V = \beta c$. This satisfies the required drift for sufficiently large β since, as we have already observed, $c(q(t))$ is a convex, decreasing function of t under an optimal policy.

The proof of (ii) is identical since the function $V(x) = \tau_\theta(x)$ is radially homogeneous and convex. ☐

Although Proposition 9 shows that optimal fluid policies are stabilizing, these policies can be highly complex, even when they are computable (again see [41, 49]). We turn next to a simpler class of policies.

The constrained-complexity optimal fluid policy. The difficulty with using an optimal state trajectory $q^*(\,\cdot\,;\,\cdot\,)$ is that complexity, as measured by the number of discontinuities in $\frac{d}{dt}q^*(t;x)$, i.e. the number of switches in the control $z^*(t;x)$, can grow exponentially with ℓ.

To bound complexity, suppose that we take a number κ and demand that q be piecewise linear, with at most κ pieces, so that the control can change no more than κ times. Any $q(\,\cdot\,;\,\cdot\,)$ which is optimal with respect to the total cost (3.1) subject to this constraint will be called a κ-constrained optimal fluid process.

Any such policy can be computed by solving a $\kappa \cdot (\ell + 1)$-dimensional quadratic program when the cost is linear [41]. The variables can be taken as the switch over times $\{0 = T_0, T_1, \ldots, T_\kappa\}$, and the control increments $\{z(T_{i+1}) - z(T_i) : 0 \le i < \kappa\}$.

Proposition 10. *Suppose that the fluid model is stabilizable. Then any κ-constrained optimal fluid process possesses a Lipschitz continuous Lyapunov function so that (A1) holds.*

Proof. One can again show that $c(q(t))$ is a convex, decreasing function of t for any κ-constrained optimal fluid process. Hence can take $V = \beta c$ for $\beta > 0$ sufficiently large. ☐

The greedy fluid policy. The greedy policy determines the allocation rate $\zeta(t)$ that minimizes $\frac{d}{dt}c(q(t))$ at each t. One motivation for this class of policies comes from considering the dynamic programing equations for the infinite-horizon optimal control problem. The optimal policy is the solution to (4.7) with c replaced by the value function V^*. Greedy heuristics are the most popular in queueing theory. The papers [8, 28] consider greedy policies for state-based cost functions as developed here. The shortest expected delay policy [32], and the least slack policy [37] are based on greedy heuristics for delay minimization.

Suppose that the cost function c is continuously differentiable (C^1). The greedy feedback law $f(x)$ is computed by solving the following ℓ-dimensional linear program:

For any x let ∇c denote the gradient of c evaluated at x, and solve

$$\begin{array}{llcll} \min\langle\nabla c, B\zeta\rangle & & & & \\ \text{subject to} & (B\zeta+\alpha)_i & \geq & 0 & \text{for all } i \text{ such that } x_i = 0 \\ & \zeta & \in & \mathsf{U}. & \end{array} \tag{4.7}$$

Then $f(x)$ is defined to be any ζ which optimizes this linear program. The linear program depends only upon $\mathrm{sign}(x)$ when the cost is linear. Given the feedback law f we then set $\frac{d}{dt}z(t;x) = f(q(t;x))$. As was seen in the previous examples, in many cases the greedy policy leads to a pathwise optimal solution - geometric conditions ensuring this are developed in [43].

The following is a generalization of a result of [10].

Proposition 11. *Suppose that the fluid model is stabilizable, and the cost function c is C^1. In this case any greedy fluid policy f is stabilizing for the fluid model, and it is pathwise optimal if a unique pathwise optimal solution exists.*

Moreover, Assumption (A1) holds with the Lyapunov function $V(x) = \beta c(x)$ for $\beta > 0$ sufficiently large.

Proof. As in the construction of a solution to (2.16), we can use stabilizability to ensure the existence of an $\varepsilon > 0$ such that the equation

$$B\zeta + \alpha = -\varepsilon\frac{x}{\|x\|}$$

has a solution $\zeta^x \in \mathsf{U}$ for any $x \neq \theta$. Hence, under the greedy policy we have, when $q(t) = x \neq \theta$,

$$\begin{aligned} \frac{d}{dt}c(q(t)) &\leq \langle\nabla c, B\zeta^x + \alpha\rangle \\ &\leq -\frac{\varepsilon}{\|x\|}\langle\nabla c, x\rangle = -\varepsilon\frac{c(x)}{\|x\|}. \end{aligned}$$

The last equality follows from radial homogeneity of the norm c. This implies the result with $\beta = \varepsilon^{-1}\max_{x\neq\theta}\{\|x\|/c(x)\}$. □

For the first four models shown above in Figures 4,

6, 8 and 10, the greedy-fluid policy is path-wise optimal. Hence it attains the minimal cost $V^*(x,T)$ for any x and any $T > 0$. The model shown in Figure 12 is a re-entrant line, for which the greedy policy is the LBFS priority policy. In this example the LBFS policy is *not* optimal for the fluid model since it results in excessive starvation at the second machine whenever the second machine is the bottleneck (see Figure 3).

Note that the greedy policy for a *discrete network* is typically defined to be the policy which, at time t, minimizes over all admissible actions a the value $\mathsf{E}[c(Q(k{+}1)) \mid Q(k), a]$. For any network scheduling problem this policy gives strict priority to exit buffers when $c(\,\cdot\,) = |\cdot|$. In general such a policy may perform extremely poorly: For the example given in Figure 10 the greedy fluid policy is path-wise optimal, but we saw in Section 2 that the priority policy is destabilizing for some parameter values even under (2.13).

4.3. Information. The policies considered thus far require the following information for successful design:

(i) The arrival rates α;

(ii) Service rates and routing information, as coded in the matrix B;

(iii) Bounds on variability of $(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{S})$ so that appropriate safety-stocks can be defined;

(iv) Global state information $q(t;x)$ for each time t.

Relaxing this information structure is of interest in various applications.

We consider here only the first issue: In telecommunications applications we may know little about arrival rates to the system, and in a manufacturing application *demand* may be uncertain. Sensitivity with respect to service and arrival rates may be large when the load is close to unity (see [19]).

To obtain a design without knowledge of arrival rates we define a set of *generalized Klimov indices* which assign priorities to buffers, subject to positivity constraints, and buffer level constraints. In this way one can define the policy in terms of observed buffer levels without knowledge of the value of α. One can address (ii) in a similar manner.

Define the permutation $(i_1, \ldots, i_\ell)$ of $\{1, \ldots, \ell\}$ so that for any resource j,

$$\langle \nabla c, Be^{i_n} \rangle \leq \langle \nabla c, Be^{i_m} \rangle \qquad \text{if } n < m, \text{ and } i_n, i_m \in \mathcal{R}_j.$$

An allocation rate $\zeta = f(x)$ can then be defined as follows: for any $n \geq 1$,

$$\sum \{\zeta_{i_k} : k \leq n, i_k \in \mathcal{R}_j\} = 1 \qquad \text{whenever} \qquad \sum \{q_{i_k} : k \leq n, i_k \in \mathcal{R}_j\} > 0. \tag{4.8}$$

Proposition 12 follows from Proposition 11 since (4.8) may be viewed as an alternative representation of the greedy allocation.

Proposition 12. *Suppose that the network is stabilizable. Then the policy (4.8) is stabilizing for the fluid model. It is pathwise optimal if a unique pathwise optimal solution exists.* □

4.4. Stability and performance. We are assured of stability under Assumptions (A1)–(A3).

Theorem 13. *Suppose that Assumptions (A1) – (A3) hold, and suppose that for some* $\varepsilon_1 \geq 0$, $\underline{N} > 0$,

(a) $\displaystyle\limsup_{\|x\|\to\infty} \frac{N(x)}{\|x\|} \leq \varepsilon_1$;

(b) $\displaystyle\limsup_{\|x\|\to\infty} \frac{\|\bar{x}(x)\|}{N(x)} \leq \varepsilon_1$;

(c) $\displaystyle\liminf_{\|x\|\to\infty} N(x) \geq \underline{N}$;

Then for all $\underline{N}$ *sufficiently large,*

(i) *The state process* $\boldsymbol{Q}$ *is ergodic in the sense that for any initial condition* $Q(0) = x$*, and any function* g *which is bounded by some polynomial function of* x*, there exists a finite* $\nu(g)$ *such that as* $T \to \infty$,

$$\frac{1}{T}\sum_{0}^{T-1} g(Q(k)) \to \nu(g) \qquad a.s.;$$
$$\mathsf{E}_x[g(Q(T))] \to \nu(g).$$

(ii) *There exists* $\Delta > 0$ *such that if* $\boldsymbol{Q}^\Delta$ *is the state process for a new network satisfying* $\|B^\Delta - B\| \leq \Delta$, $\|\alpha^\Delta - \alpha\| \leq \Delta$*, then the policy will continue to stabilize the perturbed system, in the sense of (i).*

Proof. The idea of the proof is to construct a constant b and $\varepsilon > 0$ such that

$$P^{N(x)}V(x) = \mathsf{E}[V(Q(N(x);x))] \le V(x) - \varepsilon N(x), \qquad \|x\| \ge b. \tag{4.9}$$

From the Lipschitz continuity of the model we can then find, for each p, a $b_p < \infty$ and $\varepsilon_p > 0$ such that,

$$P^{N(x)}V^p(x) \le V^p(x) - \varepsilon_p N(x) V^{p-1}(x), \qquad \|x\| \ge b_p.$$

Since V is equivalent to a norm on $\mathbb{R}^\ell$ we can argue as in [15] that

$$P^{N(x)}V^p(x) \le V^p(x) - \varepsilon_p \mathsf{E}_x\Big[\sum_0^{N(x)-1} \|Q(k)\|^{p-2}\Big], \qquad \|x\| \ge b_p,$$

where the constants b_p, ε_p may have to be adjusted, but remain finite and non-zero. It follows that the process $\mathbf{\Phi}$ is g-regular, with $g(x) = \|x\|^q$, for any $q \ge 1$. The ergodic theorems then follow, and in fact the ergodic limit in (i) converges faster than any polynomial function of time (see [15]).

How then do we establish (4.9)? For $x \in \mathsf{X}$ let $T = N(x)$ and write

$$\begin{aligned} V(Q(T;x)) &= V\Big(q(T;(x-\bar{x})^+) + (Q(T;x) - q(T;(x-\bar{x})^+)\Big) \\ &\le V(q(T;(x-\bar{x})^+) + b_0\|Q(T;x) - q(T;(x-\bar{x})^+)\|, \end{aligned}$$

where the inequality follows from Lipschitz continuity. The desired bound easily follows since $V(q(T;(x-\bar{x})^+) \le V((x-\bar{x})^+) - T$ for all x sufficiently large.

This proves (i), and (ii) follows since the drift inequality is preserved under perturbations in the model when V is Lipschitz.

Note that if N and $\bar{x}$ are bounded then, by following the proof of Theorem 4 and the arguments here, one can show that the state process $\mathbf{\Phi}$ is V_ε-uniformly ergodic.

□

Given this large class of policies, how can we compare one over another? If our goal is to estimate η, the steady state mean of $c(Q(k))$ under a given policy, and if the one step cost c is linear, then bounds on η can be obtained by solving certain linear programs [36, 12, 34], or through comparison methods with a simpler model for which performance is readily computed [48]. If these bounds are not useful then one can resort to simulation.

The standard estimator of η is given by $\widehat{\eta}(k) := k^{-1}\sum_0^{k-1} c(Q(i))$, and this estimator is strongly consistent for the policies considered here. From g-regularity with $g(\,\cdot\,) = \|\cdot\|^4$ we can also establish a central limit theorem of the form $\sqrt{t}(\widehat{\eta}(k) - \eta) \Rightarrow \sigma N(0,1)$, where $\Rightarrow$ denotes weak convergence, and $N(0,1)$ is a standard normal random variable [47]. The constant σ^2 is known as the *time-average variance constant* (TAVC), and provides a measure of the effectiveness of $\widehat{\eta}(k)$.

The problem with simulation is that the TAVC is large in heavy traffic. It is known that the TAVC is of order $(1-\rho)^{-4}$ for the M/M/1 queue [1, 57], and similar bounds hold for other network control problems [30]. With such a large variance, long run-lengths will be required to estimate η effectively.

One method of reducing variance is through control variates: For any function $h\colon \mathsf{X} \to \mathbb{R}$ let $\Delta_h = h - Ph$. Here the transition function P may define the statistics of the process $\mathbf{\Phi}$, in which case h is interpreted as a function of the first component of $\mathbf{\Phi}$ only. If the function h is π-integrable, then $\pi(\Delta_h) = 0$, and so one might use the consistent estimator

$$\widehat{\eta}_c(n) = \widehat{\eta}(n) + \frac{1}{n}\sum_{i=0}^{n-1} \Delta_h(Q(i)). \tag{4.10}$$

This approach can lead to substantial variance reductions, especially in heavy traffic, when applied to the GI/G/1 queue [29].

In [30] these ideas are extended to network models. First note that the estimator (4.10) will have a variance of *zero* when h solves the Poisson equation $\Delta_h = -c + \eta$. While this choice is not computable in general, we can approximate h by the fluid value function,

$$V(y) := \int_0^\infty c(q(t;x))\, dt.$$

The *fluid estimator* of η is then given by (4.10) with $h = V$,

$$\widehat{\eta}_f(n) = \widehat{\eta}(n) + \frac{1}{n}\sum_{k=0}^{n-1} \Delta_V(\Phi(k)). \tag{4.11}$$

Why should V provide an approximation to the solution of Poisson's equation? This actually follows from the construction of the feedback regulation policy which requires that the increments of $\boldsymbol{Q}$ approximate the increments of the fluid trajectories. See [30] for details.

REFERENCES

[1] S. Asmussen. Queueing simulation in heavy traffic. *Math. Operations Res.*, 17:84–111, 1992.

[2] N. Baëuerle. Asymptotic optimality of tracking-policies in stochastic networks. Annals of Applied Probability, To appear., 1999.

[3] S. Balaji and S.P. Meyn. Multiplicative ergodic theorems and large deviations for an irreducible Markov chain. *Stochastic Processes and their Applications*, 90(1):123–144, 2000.

[4] S.L. Bell and R.J. Williams. Dynamic scheduling of a system with two parallel servers: Asymptotic optimality of a continuous review theshold policy in heavy traffic. In *Proceedings of the 38th Conference on Decision and Control*, pages 1743–1748, Pheonix, Arizona, 1999.

[5] D. Bertsekas and R. Gallager. *Data Networks: Upper Saddle River*. Prentice-Hall, Englewood Cliffs, NJ, 1992.

[6] M. Bramson. Instability of FIFO queueing networks. To appear in *Annals of Applied Probability*, 1994.

[7] H. Chen and A. Mandelbaum. Discrete flow networks: Bottlenecks analysis and fluid approximations. *Mathematics of Operations Research*, 16:408–446, 1991.

[8] H. Chen and D. D. Yao. Dynamic scheduling of a multiclass fluid network. *Operations Research*, 41(6):1104–1115, November–December 1993.

[9] R-R. Chen and S.P. Meyn. Value iteration and optimization of multiclass queueing networks. *Queueing Systems*, 32:65–97, 1999. Invited Paper.

[10] D. P. Connors, G. Feigin, and D. Yao. Scheduling semiconductor lines using a fluid network model. *IEEE Trans. on Robotics and Automation*, 10:88–98, 1994.

[11] R. L. Cruz. A calculus for network delay, part I: Network elements in isolation. *IEEE Trans. Inform. Theory*, 37(1):114–131, January 1991.

[12] I. Paschalidis D. Bertsimas and J. N. Tsitsiklis. Optimization of multiclass queueing networks: polyhedral and nonlinear characterizations of achievable performance. *Ann. Appl. Probab.*, 4:43–75, 1994.

[13] J. Humphrey D. Eng and S.P. Meyn. Fluid network models: Linear programs for control and performance bounds. In J. Cruz J. Gertler and M. Peshkin, editors, *Proceedings of the 13th IFAC World Congress*, volume B, pages 19–24, San Francisco, California, 1996.

[14] J. G. Dai. On the positive Harris recurrence for multiclass queueing networks: A unified approach via fluid limit models. *Ann. Appl. Probab.*, 5:49–77, 1995.

[15] J. G. Dai and S.P. Meyn. Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Trans. Automat. Control*, 40:1889–1904, November 1995.

[16] J. G. Dai and G. E. Weiss. Stability and instability of fluid models for reentrant lines. *Math. Operations Res.*, 21:115–134, 1996.

[17] J. G. Dai and G. E. Weiss. A fluid heuristic for minimizing makespan in job-shops. Preprint, Georgia Inst. of Technology, 1999.

[18] B. T. Doshi. Optimal control of the service rate in an M/G/1 queueing system. *Adv. Appl. Probab.*, 10:682–701, 1978.

[19] P. Dupuis and K. Ramanan. A multiclass feedback queueing network with a regular skorokhod problem. *Preprint*, Brown University, 1999.

[20] P. Dupuis and R. J. Williams. Lyapunov functions for semimartingale reflecting Brownian motions. *Ann. Appl. Probab.*, 22(2):680–702, 1994.

[21] S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1:397–413, 1993.

[22] S. B. Gershwin. *Manufacturing Systems Engineering*. Prentice–Hall, Englewood Cliffs, NJ, 1993.

[23] B. Hajek. Optimal control of two interacting service stations. *IEEE Trans. Automat. Control*, AC-29:491–499, 1984.

[24] J. M. Harrison. Brownian models of heterogeneous customer populations. In W. Fleming and P. L. Lions, editors, *Stochastic Differential Systems, Stochastic Control Theory and Applications, IMA Volumes in Mathematics and its Applications*, chapter 10, pages 147–186. Springer-Verlag, 1988.

[25] J. M. Harrison. Brownian models of open processing networks: Canonical representations of workload. *Ann. Appl. Probab.*, 10:75–103, 2000.

[26] J. M. Harrison and M.J. López. Heavy traffic resource pooling in parallel-server systems. Preprint, Stanford University, 1998.

[27] J. M. Harrison and L. M. Wein. Scheduling networks of queues: Heavy traffic analysis of a simple open network. *QUESTA*, 5:265–280, 1989.

[28] J.M. Harrison. *The BIGSTEP approach to flow management in stochastic processing networks*, pages 57–89. Stochastic Networks Theory and Applications. Clarendon Press, Oxford, UK, 1996. F.P. Kelly, S. Zachary, and I. Ziedins (ed.).

[29] S.G. Henderson. *Variance Reduction Via an Approximating Markov Process*. PhD thesis, Stanford University, Stanford, California, USA, 1997.

[30] S.G. Henderson and S.P. Meyn. Variance reduction for simulation in multiclass queueing networks. submitted to the IIE Transactions on Operations Engineering: special issue honouring Alan Pritsker on simulation in industrial engineering, 1999.

[31] C. Humes, Jr., J. Ou, and P. R. Kumar. The delay of open Markovian queueing networks: Uniform functional bounds, heavy traffic pole multiplicities, and stability. *Mathematics of Operations Research*, 22(4):921–954, November 1997.

[32] F.C. Kelly and C.N. Laws. Dynamic routing in open queueing networks: brownian models, cut constraints and resource pooling. *QUESTA*, 13:47–86, 1993.

[33] Y. Kontoyiannis and S. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. Submitted for publication, 2000.

[34] P.R. Kumar and S.P. Meyn. Duality and linear programs for stability and performance analysis queueing networks and scheduling policies. *IEEE Transactions on Automatic Control*, 41(1):4–17, January 1996.

[35] P.R. Kumar and T. I. Seidman. Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems. *IEEE Trans. Automat. Control*, AC-35(3):289–298, March 1990.

[36] S. Kumar and P. R. Kumar. Performance bounds for queueing networks and scheduling policies. *IEEE Trans. Automat. Control*, AC-39:1600–1611, August 1994.

[37] S. Kumar and P. R. Kumar. Fluctuation smoothing policies are stable for stochastic re-entrant lines. *Journal of Discrete Event Dynamic Systems: Theory and Applications*, 6(4):361–370, October 1996.

[38] S. Kumar and M. Muthuraman. A numerical method for solving singular brownian control problems. In *Proceedings of the 39th Conference on Decision and Control*, 2000.

[39] W. Lin and P. R. Kumar. Optimal control of a queueing system with two heterogeneous servers. *IEEE Trans. Automat. Control*, AC-29:696–703, August 1984.

[40] S. Lippman. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 23:687–710, 1975.

[41] X. Luo and D. Bertsimas. A new algorithm for state-constrained separated continuous linear programs. *SIAM J. Control Optim.*, 37:177–210, 1998.

[42] C. Maglaras. Design of dynamic control policies for stochastic processing networks via fluid models. In *Proceedings of the 38th Conference on Decision and Control*, pages 1208–1213, 1997.

[43] S. P. Meyn. Sequencing and routing in multiclass queueing networks. Part II: Workload relaxations. Submitted for publication, 2000.

[44] S.P. Meyn. The policy improvement algorithm for Markov decision processes with general state space. *IEEE Trans. Automat. Control*, AC-42:1663–1680, 1997.

[45] S.P. Meyn. Stability and optimization of multiclass queueing networks and their fluid models. In *Mathematics of Stochastic Manufacturing Systems. AMS-SIAM Summer Seminar in Applied Mathematics*, volume 33 of *Lectures in Applied Mathematics*, pages 175–199. American Mathematical Society, Providence, RI, 1997.

[46] S.P. Meyn and D. Down. Stability of generalized Jackson networks. *Ann. Appl. Probab.*, 4:124–148, 1994.

[47] S.P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993.

[48] J. Ou and L. M. Wein. Performance bounds for scheduling queueing networks. *Ann. Appl. Probab.*, 2:460–480, 1992.

[49] J. R. Perkins and P. R. Kumar. Optimal control of pull manufacturing systems. *IEEE Trans. Automat. Control*, 40(12):2040–2051, December 1995.

[50] M.C. Pullan. Forms of optimal solutions for separated continuous linear programs. *SIAM J. Control Optim.*, 33:1952–1977, 1995.

[51] C. Okino R. Agrawal, R.L. Cruz and R. Rajan. Performance bounds for flow control protocols. *IEEE Trans. on Networking*, 7:310–323, 1999.

[52] A. N. Rybko and A. L. Stolyar. On the ergodicity of stohastic processes describing the operation of open queueing networks. *Problemy Peredachi Informatsii*, 28:3–26, 1992.

[53] J. G. Shanthikumar and D. D. Yao. Stochastic monotonicity in general queueing networks. *J. Appl. Probab.*, 26:413–417, 1989.

[54] A. Shwartz and A. Weiss. *Large deviations for performance analysis: queues, communication and computing*. Chapman-Hall, London, UK, 1995.

[55] Stephen Lavenberg, Editor. *Computer Performance Modeling Handbook*. Academic Press, New York, 1983.

[56] G. Weiss. On the optimal draining of re-entrant fluid lines. Technical report, Georgia Institute of Technology and Technion, 1994.

[57] W. Whitt. Planning queueing simulations. *Management Science*, 35:1341–1366, 1994.

[58] 6.S.H. Xu and H. Chen. A note on the optimal control of two interacting service stations. *IEEE Transactions on Automatic Control*, 38:187–189, 1993.