Dynamic safety-stocks for asymptotic optimality in stochastic networks*

Sean Meyn†

May 10, 2005

Abstract

This paper concerns control of stochastic networks using state-dependent safety-stocks. Three examples are considered: a pair of tandem queues; a simple routing model; and the Dai-Wang re-entrant line. In each case, a single policy is proposed that is independent of network load $\rho_\bullet$. The following conclusions are obtained for the controlled network in each case, where the finite constant K_0 is independent of load: The policy is fluid-scale asymptotically optimal, and approximately average-cost optimal. The steady-state cost η satisfies the bound

$$\eta_* \le \eta \le \eta_* + K_0 \log(\eta_*), \qquad 0 < \rho_\bullet < 1,$$

where η_* is the optimal steady-state cost.

These results are based on the construction of an approximate solution to the average-cost dynamic programming equations using a perturbation of the value function for an associated fluid model.

Moreover, a new technique is introduced to obtain fluid-scale asymptotic optimality for general networks modeled in discrete time. The proposed policy is myopic with respect to a perturbation of the value function for the fluid model.

Keywords: Queueing Networks, Routing, Scheduling, Optimal Control.

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

*This paper is based upon work supported by the National Science Foundation under Award Nos. ECS 02 17836 and DMI 00 85165.

†Department of Electrical and Computer Engineering and the Coordinated Sciences Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A. (meyn@uiuc.edu).

1 Introduction

Consider the simple routing model shown at left in Figure 1. Customer arrivals to the system are modeled as a Poisson process with rate α_1, and are routed to one of the two servers based on a particular policy. Service times at the two downstream stations are i.i.d. and exponentially distributed. Upon sampling, this system can be modeled as a discrete time Markov decision process (MDP). Hence an optimal policy can be obtained using dynamic programming methods.

In this paper we consider general ℓ-dimensional MDP network models along with their deterministic 'fluid' analogs. The following two control criteria for the stochastic model are based on a given cost function $c\colon \mathbb{R}_+^\ell \to \mathbb{R}_+$. The *average-cost* optimality criterion concerns the steady-state behavior of the network; and the *fluid-scale* optimality criterion is based on the transient behavior from large initial conditions. The latter view of optimization has received significant attention recently due to its tractability, and the resulting stability properties that can be obtained for the controlled network [10, 28, 26, 39].

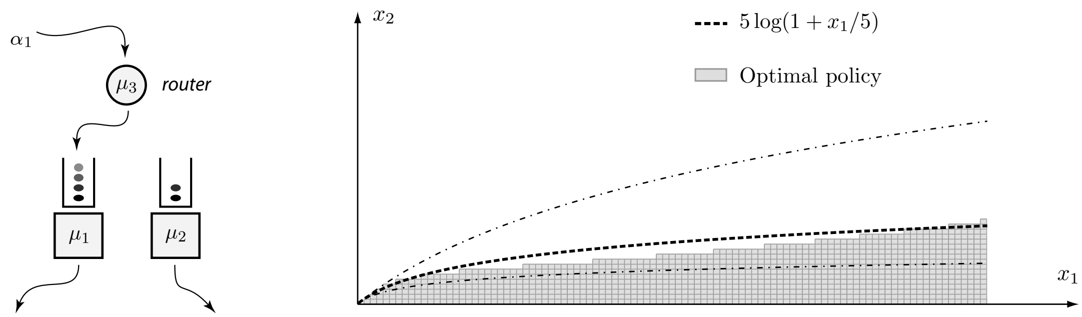


Figure 1: The optimal policy for the simple routing model. The grey region in the plot at right shows the optimal switching curve obtained using value iteration. Within this region, arriving packets are routed to buffer 2. The network parameters are $\alpha = 9/19$, $\mu_1 = \mu_2 = 5/19$, and the one step cost is $c(x) = 2x_1 + 3x_2$. The three concave curves are plots of the switching curve $s_\gamma(x_1) = \gamma \log(1 + x_1/\gamma)$ for $\gamma = 2, 5, 20$. The best fit is when $\gamma = 5$. This closely matches the lower bound presented in Proposition 4.1, which in this instance is equal to $3\bar{\beta}^{-1} = 3/\log(2) \approx 5.1$.

Optimality under each of these criteria is formally defined as follows: A policy is called *average-cost optimal* for the stochastic model if the average cost,

$$\eta := \limsup_{t\to\infty} \mathsf{E}[c(Q(t;x))] \tag{1}$$

is minimized over all policies, where $Q(t;x)$ denotes the vector of buffer-lengths at time t starting from the initial condition $x \in \mathbb{N}^\ell$. The average cost (1) is independent of x under the optimal policy for the models considered in this paper [26].

The ℓ-dimensional queue-length process $Q(t;x)$ is defined in (12) below in discrete-time, but we may interpolate to obtain a piecewise linear function of $t \in \mathbb{R}_+$. We then define for each $n \geq 1$, $x \in \mathbb{R}_+^\ell$,

$$\begin{aligned} q_n(t;x) &:= n^{-1} Q(nt;[nx]), \qquad t \in \mathbb{R}_+, \\ J_n(x,T) &:= \mathsf{E}\Big[\int_0^T c(q_n(t;x))\,dt\Big], \qquad T \in \mathbb{R}_+, \end{aligned} \tag{2}$$

where $[y] \in \mathbb{N}^\ell$ denotes the integer part of $y \in \mathbb{R}^\ell_+$. A policy is called *fluid-scale asymptotically optimal* (FSAO) if

$$\limsup_{n\to\infty} J_n(x,T) \le J_*(x), \qquad x \in \mathbb{R}^n_+,\ T \ge 0, \tag{3}$$

where J_* denotes the optimal value function for the fluid model (see (7) below.) For the class of network models considered in this paper, it follows as in [26, Theorem 7.2] that the average-cost optimal policy for the stochastic model is always FSAO (see also [1].)

Safety-stocks are commonly used to prevent starvation of resources for effective control of a stochastic network (see e.g. [33, 24, 2, 4, 37].) Several recent papers explore the application of switching curves that grow logarithmically with increasing network congestion to define *dynamic* safety-stocks. A general approach of this form was introduced in [27, p. 194] based on numerical results obtained for several examples including the routing model shown in Figure 1 and the 'processor sharing model' of [13, 16, 2, 32]. Further motivation is provided in [29, p. 207] based on a 'heavy-traffic' analysis. A policy of this form is analysed in depth in [10] for a pair of queues in tandem. It is found that the policy is stabilizing, and FSAO. Moreover, a bound is obtained on the rate of convergence in (3).

Structure of optimal policies for routing models of the form illustrated in Figure 1 is the subject of [14], and [20, 37] each contain analyses of threshold policies for approximate optimality. More closely related to the research reported here is [38], in which optimal policies are obtained for both stochastic and fluid routing models. The optimal policy for the particular model shown above is illustrated at right in Figure 1, along with several instances of the logarithmic switching curves considered in this paper. In this example it is apparent that the optimal policy can be closely approximated by a switching curve of this simple logarithmic form.

In each of the examples considered in this paper it is shown that one can construct a fixed policy based on a logarithmic switching curve that is FSAO when the load conditions hold, and approximately average-cost optimal as the load increases to unity. The proof is based on the construction of an approximate solution to the following average-cost dynamic programming equations for the stochastic network: the *relative value function* $h_*\colon \mathbb{R}^\ell_+ \to \mathbb{R}$ and constant η_* jointly satisfy,

$$\min_u \mathsf{E}[h_*(Q(k+1)) - h_*(Q(k)) \mid Q(k)=x, U(k)=u] = -c(x) + \eta_*, \qquad x \in \mathbb{N}^\ell, \tag{4}$$

where $Q(k)$ is the vector of queue-lengths, and $U(k)$ is the vector of allocation values at time k. The minimum in (4) is over all allowable allocation values. Under a quadratic growth condition on h_* and second moment assumptions on the model, the dynamic programming equation implies that η_* is the minimum of the average cost (1) over all policies [26].

Some of the conclusions in this paper are generalizations of well-known structural properties of the following simple queueing model. This is a version of the *controlled random walk* model that is considered in various multi-dimensional settings below.

For an initial condition $Q(0) = x \in \mathbb{N}$, the queue-length process in the simple queue evolves according to the recursion,

$$Q(k+1) = Q(k) - S(k+1)U(k) + A(k+1), \qquad k \ge 0. \tag{5}$$

It is assumed that the joint arrival-service process $\{A(k), S(k) : k \ge 1\}$ is i.i.d. with a finite second moment; the common marginal distribution of $\boldsymbol{S} = \{S(k) : k \ge 1\}$ is binary; and the marginal distribution of $\boldsymbol{A}$ is supported on $\mathbb{N}$. When the allocation sequence $\boldsymbol{U}$ is defined as the non-idling policy (i.e. $U(k) = 1$ if $Q(k) \ge 1$), then $\boldsymbol{Q}$ is a Markov chain on $\mathbb{N}$.

It is well-known that upon sampling via uniformization, an M/M/1 queue can be represented as the solution to the recursion (5) using a non-idling policy. In this special case, the arrival stream is also binary valued, with $\alpha = \mathsf{P}(A(k) = 1) = 1 - \mathsf{P}(S(k) = 1) = 1 - \mu$.

The associated fluid model is given by the ordinary differential equation (ODE) model,

$$\frac{d^+}{dt} q(t) = -\mu\zeta(t) + \alpha\,, \qquad t \ge 0,$$

where the '+' denotes right derivative, and $\alpha = \mathsf{E}[A(1)]$, $\mu = \mathsf{E}[S(1)]$. The non-idling policy is defined by $\zeta(t) = 1$ if $q(t) > 0$, and $\zeta(t) = \rho = \alpha/\mu$ if $q(t) = 0$. On taking the cost function to be $c(x) = x$, the fluid value function under the non-idling policy is

$$J_*(x) := \int_0^\infty q(t)\,dt = \frac{1}{2}\frac{1}{\mu - \alpha}x^2, \qquad q(0) = x \in \mathbb{R}_+. \tag{6}$$

The following proposition is required in our treatment of one-dimensional workload relaxations, and is used to illustrate our main conclusions. Observe that the steady state mean η_* given in Proposition 1.1 (i) is analogous to the steady state mean of the GI/M/1 queue [19, 21] and also the reflected Brownian motion [4, Theorem 6.2]. A proof is provided in the Appendix.

Proposition 1.1. *Consider the general one-dimensional queueing model (5) under the non-idling policy. Suppose that* $\rho = \alpha/\mu < 1$ *and* $\sigma^2 := \rho m^2 + (1-\rho)m_A^2 < \infty$*, where* $m^2 = \mathsf{E}[(S(1) - A(1))^2] < \infty$, $m_A^2 = \mathsf{E}[A(1)^2] < \infty$*. Then,*

(i) *There is a unique invariant probability distribution* π *on* $\mathbb{N}$ *satisfying*

$$\eta_* := \mathsf{E}_\pi[Q(k)] = \frac{1}{2}\frac{\sigma^2}{\mu - \alpha}, \qquad k \ge 0.$$

(ii) *The function* $h_* : \mathbb{N} \to \mathbb{R}_+$ *defined as*

$$h_*(x) := J_*(x) + \frac{1}{2\mu}\Big(\frac{m^2 - m_A^2}{\mu - \alpha}\Big)x\,, \qquad x \in \mathbb{N},$$

solves Poisson's equation, $\mathsf{E}[h_*(Q(k+1)) - h_*(Q(k)) \mid Q(k) = x] = -c(x) + \eta_*$. ☐

It is assumed in the main result Theorem 2.2 that there is a single bottleneck in heavy traffic. Under this restriction it is possible to construct a lower bound on the average cost (1) that is shown to be tight as the network load approaches unity. This bound was first introduced in the thesis [23] of N. Laws: A one-dimensional workload process $\boldsymbol{W}$ corresponding to the most heavily loaded station is compared to its relaxation $\widehat{\boldsymbol{W}}$, which is a version of the simple queueing model (5). The two processes are defined using the same driving sequences $\boldsymbol{A}$ and $\boldsymbol{S}$. For each $k \ge 0$ the lower bound $W(k) \ge \widehat{W}^*(k)$ holds with probability one, under *any* policy for $\boldsymbol{Q}$, where the relaxation $\widehat{\boldsymbol{W}}^*$ is controlled using the non-idling policy.

We restrict to models of a generalized 'Kelly type' in which service statistics are determined by the station, not the buffer. This assumption is imposed to facilitate the construction of a workload relaxation (see [5, 29] for more details on workload relaxations, and [18, 23] for related techniques.)

The approximation of the relative value function for the multidimensional network is based on the value function J_* for the fluid model. The fluid model evolves on $\mathbb{R}^\ell_+$, with state denoted $q(t)$ at time t. The value function J_* is defined by

$$J_*(x) := \min \int_0^\infty c(q(t))\,dt, \quad q(0) = x \in \mathbb{R}^\ell_+, \tag{7}$$

where the minimum is over all policies. In each example the cost is assumed linear, and it is found that (i) J_* is a perturbation of the relative value function for the one-dimensional workload relaxation $\widehat{\boldsymbol{W}}^*$, and (ii) J_* *almost* solves the average-cost dynamic programming equations for the ℓ-dimensional network. These conclusions follow in part from the following dynamic programming equation for the fluid model: In the examples considered in this paper the fluid value function is convex, C^1, and satisfies,

$$\min_v \langle \nabla J_*(x), v\rangle = \langle \nabla J_*(x), v^*(x)\rangle = -c(x), \qquad x \in \mathbb{R}^\ell_+, \tag{8}$$

where $v_*(x) = \frac{d^+}{dt} q_*(t)$ when $q(t) = x$, and the minimum is over all feasible velocity vectors. Moreover, the gradient ∇J_* is Lipschitz continuous. Consequently, the following bound follows from the Mean Value Theorem for the discrete-time stochastic model,

$$\mathsf{E}[J_*(Q(k+1)) - J_*(Q(k)) \mid Q(k) = x,\ U(k) = u] \le \langle \nabla J_*(x), v^0(x)\rangle + \mathcal{E}_0(x) \tag{9}$$

where $v^0(x) := \mathsf{E}[Q(k+1) - Q(k) \mid Q(k) = x,\ U(k) = u]$, and the following error term is bounded,

$$\mathcal{E}_0(x) := \mathsf{E}\big[\big|\langle \nabla J_*\left(Q(k+1)\right) - \nabla J_*\left(Q(k)\right), Q(k+1) - Q(k)\rangle\big| \mid Q(k) = x,\ U(k) = u\big]$$

The inner-product on the right hand side of (9) is equal to $-c(x)$, provided that the policy for the stochastic model is consistent with the optimal policy for the fluid model in the sense that $v^0(x) = v_*(x)$. This is feasible, possibly using a randomized policy, provided $x_i \ge 1$ for each i.

The identity $v^0(x) = v_*(x)$ may be infeasible along the boundaries of the state space. An additive correction term is introduced to account for these inconsistencies: We define

$$V(x) = J_*(x) + b(x), \qquad x \in \mathbb{N}^\ell, \tag{10}$$

where b has at most linear growth. In each example considered the correction term b is constructed so that the function V satisfies a *Lyapunov drift condition* under the proposed policy, of the form,

$$\mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x] \le -c(x) + \widehat{\eta}_* + \mathcal{E}(x), \qquad x \in \mathbb{N}^\ell, \tag{11}$$

where $\widehat{\eta}_*$ is the optimal cost for the relaxation $\widehat{\boldsymbol{W}}^*$, and the error $\mathcal{E}$ has at most logarithmic growth. This inequality is a relaxation of the dynamic programming equation (4).

Based on the bound (11) it is shown that for a finite constant K_0 that is independent of load,

$$\widehat{\eta}_* \le \eta \le \widehat{\eta}_* + K_0 \log(\widehat{\eta}_*), \qquad 0 < \rho_\bullet < 1.$$

The average-cost $\widehat{\eta}_*$ is unbounded as the network load $\rho_\bullet$ approaches unity, and is of order $(1-\rho_\bullet)^{-1}$ by Proposition 1.1.

A closely related form of asymptotic optimality was first formulated in [29, Section 4.2], and verified for a class of multiclass networks with multiple bottlenecks and renewal inputs. The analysis of [29] is based on large-deviation estimates to bound idleness, following [2]. The present paper is the first to obtain such tight bounds on the average cost, and the first to establish asymptotic optimality with respect to a linear cost function using a fixed policy independent of load.

The results obtained here are also related to the approach of [15, 22] where the controlled process is compared with a reflected Brownian motion for $\rho_\bullet \sim 1$. While a useful approach for verification of a given family of policies, these methods do not currently provide any way to obtain tight bounds on the steady-state cost for $\rho_\bullet < 1$.

The remainder of the paper is organized as follows. The next section provides a description of the models considered, and a central result on network optimization, Theorem 2.2, that identifies sufficient conditions for heavy-traffic and fluid-scale asymptotic optimality. Sections 3 and 4 contain detailed analyses of two examples, based on Theorem 2.2. Section 5 contains numerical results for a more complex multiclass network.

A completely general approach to policy synthesis is presented in Section 6. The proposed policy is myopic with respect to a Lyapunov function satisfying a bound similar to (11). The function V is obtained using the fluid value function J_* combined with a state transformation. Stability and FSAO are established for this policy in Theorem 6.3.

Conclusions are contained in Section 7.

2 Controlled random walk model and its relaxations

This section contains a description of the models considered in this paper to describe a multiclass queueing network, following [28, 5, 29, 17]. We also present a Lyapunov criterion for fluid-scale and heavy-traffic asymptotic optimality.

In the controlled random walk (CRW) network model the queue length process evolves according to the recursion,

$$Q(k+1) = Q(k) + B(k+1)U(k) + A(k+1), \qquad k \geq 0,\ Q(0) = x\,. \tag{12}$$

There are ℓ buffers in the network so that the queue length process $\boldsymbol{Q}$ evolves on $\mathbb{N}^\ell$. The allocation sequence $\boldsymbol{U}$ and the arrival sequence $\boldsymbol{A}$ are also ℓ-dimensional, and $\boldsymbol{B}$ is an $\ell \times \ell$ matrix seqeunce.

There are ℓ_r stations: the $\ell_r \times \ell$ *constituency matrix*, denoted C, has binary entries, with $C_{ji} = 1$ if buffer i resides at station j, and zero otherwise. The allocation vector $U(k) \in \mathbb{N}^\ell$ is assumed to have binary entries. If $U_i(k) = 1$, then buffer i receives priority at its respective station at time k.

We let R denote the $\ell \times \ell$ routing matrix defined as $R_{i_1 i_2} = 1$ if and only if jobs leaving buffer i_1 then move to buffer i_2. The networks considered in this paper are assumed to be *open*, so that any customer entering the system can eventually exit. This is equivalent to the assumption that $R^N = 0$ for some integer $N < \ell$.

The following specifications for the matrix sequence $\boldsymbol{B}$ are based on the assumption that service statistics are determined by the station, not the buffer: An associated *service process* is denoted $S(k) = [S_1(k), \ldots, S_{\ell_r}(k)]^{\mathrm{T}}$, where $S_j(k)$ takes on binary values for $j = 1, \ldots, \ell_r$. It is assumed that $B(k) = -[I - R^{\mathrm{T}}]M(k)$, where $\boldsymbol{M}$ denotes the the diagonal matrix sequence,

$$M(k) = \mathrm{diag}(S(k)^{\mathrm{T}}C), \quad k \geq 1.$$

That is, $M_{ii}(k) = S_j(k)$ if and only if $C_{ji} = 1$.

The following additional assumptions are imposed on the policy and parameters:

(A1) The sequence $\{\binom{A(k)}{S(k)} : k \geq 1\}$ is an i.i.d. sequence of $(\ell + \ell_r)$-dimensional vectors. For each $k \geq 1$, $1 \leq j \leq \ell_r$, the distribution of $S_j(k)$ is binary, and satisfies,

$$\mathsf{P}\{S_j(k) = 1 \text{ and } A(k) = 0\} > 0. \tag{13}$$

The distribution of $A(k)$ is supported on $\mathbb{N}^\ell$, and satisfies $\mathsf{E}[\|A(k)\|^2] < \infty$.

(A2) The allocation sequence $\boldsymbol{U}$ satisfies $U(k) \in \mathsf{U}_0$ for each $k \geq 0$, where

$$\mathsf{U}_0 = \{u \in \{0,1\}^\ell : Cu \leq 1\},$$

1 denotes the vector consisting entirely of ones, and inequalities between vectors are interpreted component-wise. Moreover, $\boldsymbol{U}$ is defined by a (possibly randomized) stationary policy: there is a Borel measurable function $f\colon \mathbb{N}^\ell \times \mathbb{R}^\ell \to \mathsf{U}_0$, together with an i.i.d. stochastic process $\boldsymbol{\Psi}$ on $\mathbb{R}^\ell$ that is independent of $(Q(0), \boldsymbol{A}, \boldsymbol{S})$, such that

$$U(k) = f(Q(k), \Psi(k)), \qquad k \geq 0. \tag{14}$$

(A3) The queue length process $\boldsymbol{Q}$ is constrained to $\mathbb{N}^\ell$. Hence the allocation sequence is subject to the state-dependent constraint $U(k) \in \mathsf{U}_0(x)$ when $Q(k) = x$, where

$$\mathsf{U}_0(x) = \{u \in \mathsf{U}_0 : u_i = 0 \text{ when } x_i = 0\}. \tag{15}$$

When controlled using a stationary policy, randomized or not, the resulting stochastic process $\boldsymbol{Q}$ is a Markov chain on $\mathbb{N}^\ell$. The lower bound (13) is imposed to ensure that for the stationary policies considered, the resulting Markov chain is 0-*irreducible* in the sense that,

$$\sum_{k=1}^{\infty} \mathsf{P}_x\{Q(k) = 0\} > 0, \qquad x \in \mathbb{N}^\ell. \tag{16}$$

We denote the common mean of the arrival and service variables using the usual notation,

$$\mathsf{E}[A_i(k)] = \alpha_i, \quad \mathsf{E}[S_j(k)] = \mu_j, \qquad 1 \leq i \leq \ell,\ 1 \leq j \leq \ell_r,$$

and we let $B = \mathsf{E}[B(k)]$. The load vector is expressed,

$$\rho = -CB^{-1}\alpha, \tag{17}$$

and the system load is the maximum, $\rho_\bullet = \max_j \rho_j$. Station j is called a bottleneck if it attains this maximum.

In this paper we consider one dimensional relaxations based on the following $\ell_r \times \ell$ *workload matrix*,

$$\Xi = C[I - R^{\mathrm{T}}]^{-1}.$$

The inverse exists as a finite power series since the network is assumed open. The workload process associated with the CRW model is defined as $W(k) = C[I - R^{\mathrm{T}}]^{-1}Q(k)$, $k \ge 0$. For each $1 \le j \le \ell_r$, and time $k \ge 0$, the quantity $W_j(k) \in \mathbb{N}$ is interpreted as the workload at station j in *units of customers*.

For each $1 \le j \le \ell_r$ we let $\xi^j \in \mathbb{N}^\ell$ denote the jth row of Ξ. The jth component of the workload process is expressed $W_j(k) = \langle \xi^j, Q(k)\rangle$, and evolves according to the recursion,

$$W_j(k+1) = W_j(k) - S_j(k+1) + S_j(k+1)I_j(k) + L_j(k+1), \qquad k \ge 0, \tag{18}$$

where $L_j(k) = \langle \xi^j, A(k)\rangle$, and the idleness process $\boldsymbol{I}_j$ is defined by

$$I_j(k) = 1 - \sum_{i:C_{ji}=1} U_i(k).$$

Note that $I_j(k)$ always refers to the *noncumulative* idleness at station j, and satisfies $0 \le I_j(k) \le 1$ for each k.

We let λ_j denote the common mean of $\{L_j(k) : k \ge 1\}$, interpreted as the "mean effective arrival rate" to the jth station. The jth load parameter defined in (17) can be expressed $\rho_j = \lambda_j/\mu_j$ for $1 \le j \le \ell_r$.

The one-dimensional relaxation is defined on the same probability space with $\boldsymbol{Q}$, and evolves as a controlled random walk on $\mathbb{N}$,

$$\widehat{W}_j(k+1) = \widehat{W}_j(k) - S_j(k+1) + S_j(k+1)\hat{I}_j(k) + L_j(k+1), \qquad k \ge 0. \tag{19}$$

This is identical in form to the previous recursion, but in this relaxation we allow the idleness process $\hat{\boldsymbol{I}}_j$ to take any value in $\mathbb{N}$. It is assumed that the idleness process is adapted to $\boldsymbol{A}, \boldsymbol{S}$, with non-negative integer values, and that $\widehat{\boldsymbol{W}}_j$ is constrained to $\mathbb{N}$. On optimizing, we will choose $\hat{\boldsymbol{I}}_j$ equal to the non-idling policy for (19).

The ℓ-dimensional fluid model evolves in continuous time, and is defined using the mean-parameters for the CRW model,

$$q(t) = x + Bz(t) + \alpha t, \qquad t \ge 0, \qquad q(0) = x. \tag{20}$$

The deterministic cumulative allocation process $\boldsymbol{z}$ is subject to the constraints

$$z(t) - z(s) \ge 0, \quad C[z(t) - z(s)] \le (t-s)\mathbf{1}, \qquad 0 \le s \le t.$$

A relaxation for the fluid model is defined as in (19): For a given $j \in \{1, \dots, \ell_r\}$, the workload process is respresented as the solution to the controlled ODE,

$$\frac{d^+}{dt}\widehat{w}_j(t) = -(\mu_j - \lambda_j) + \iota(t), \qquad t \ge 0,$$

where $\widehat{\boldsymbol{w}}_j$ is constrained to $\mathbb{R}_+$, and the idleness rate $\boldsymbol{\iota}$ is non-negative.

Given a cost function $c\colon \mathbb{R}_+^\ell \to \mathbb{R}_+$ for the fluid or stochastic ℓ-dimensional model, a cost function known as the *effective cost* $\overline{c}\colon \mathbb{R}_+ \to \mathbb{R}$ is defined for the one-dimensional relaxation as the solution to the following non-linear program,

$$\begin{array}{rlrcl} \overline{c}(w) = & \min \ c(x) & \text{s.t.}\ \langle \xi^j, x\rangle & = & w \\ & & x & \in & \mathbb{R}_+^\ell . \end{array} \tag{21}$$

For a given $w \in \mathbb{R}_+$, the *effective state* $x^*(w)$ is defined to be any $x \in \mathbb{R}_+^\ell$ that optimizes (21).

In the following two typical situations the effective cost is easily computed:

(i) *Linear cost function* If $c(x) = c^{\mathrm{T}}x$, $x \in \mathbb{R}_+^\ell$, then the effective state is given by

$$x^*(w) = \Bigl(\frac{1}{\xi^j_{i^*}} e^{i^*}\Bigr) w, \qquad w \in \mathbb{R}_+,$$

where the index i^* is any solution to $c_{i^*}/\xi^j_{i^*} = \min_{1\le i\le \ell}\bigl(c_i/\xi^j_i\bigr)$. The effective cost is thus given by the linear function, $\overline{c}(w) = c(x^*(w)) = (c_{i^*}/\xi^j_{i^*})w$, $w \in \mathbb{R}_+$.

(ii) *Quadratic cost function* If $c(x) = \frac{1}{2}x^{\mathrm{T}}Dx$, $x \in \mathbb{R}_+^\ell$, with $D > 0$ then the effective state is again linear in the workload value. We have the explicit expression,

$$x^*(w) = \Bigl(\bigl({\xi^j}^{\mathrm{T}} D^{-1}\xi^j\bigr)^{-1} D^{-1}\xi^j\Bigr) w, \qquad w \in \mathbb{R}_+,$$

provided $D^{-1}\xi^j \in \mathbb{R}_+^\ell$ so that x^* is feasible. In this case the effective cost is the one-dimensional quadratic,

$$\overline{c}(w) = \tfrac{1}{2}\bigl({\xi^j}^{\mathrm{T}} D^{-1}\xi^j\bigr)^{-1} w^2, \qquad w \in \mathbb{R}_+.$$

The feasibility constraint is always satisfied when D is a positive-definite diagonal matrix.

In cases (i) and (ii) the effective cost is monotone. Consequently, we arrive at the following bound for the stochastic and fluid models,

$$\begin{aligned} c(Q(k)) &\ \ge\ \overline{c}(\widehat{W}^*_j(k)), \qquad & k \in \mathbb{N}, \\ c(q(t)) &\ \ge\ \overline{c}(\widehat{w}^*_j(t)), \qquad & t \in \mathbb{R}_+, \end{aligned} \tag{22}$$

where each process has the common initialization, with $\widehat{W}^*_j(0) = \widehat{w}^*_j(0) = \langle \xi^j, Q(0)\rangle$. The 'star' is used in this notation to stress that the relaxed model is controlled using the non-idling policy.

When the cost function c is linear one can construct a solution for the fluid model that couples with $\widehat{\boldsymbol{w}}^*$ in the sense that $q(t) = (\xi^j_{i^*})^{-1} e^{i^*}\langle \xi^j, q(t)\rangle$ after a transient period T_1 that is bounded as $\rho_j \uparrow 1$ (see [29, Theorem 4.1]). A similar result can be obtained when c is quadratic. Consequently, one has $c(q(t)) = \overline{c}(\widehat{w}^*(t))$ for $t \ge T_1$. This one-dimensional state-space collapse is also seen in heavy-traffic limit theory of stochastic networks [18, 3, 40, 29].

Observe that in the case of quadratic cost with $D > 0$ diagonal, then $x^*_i > 0$ for each buffer i at station j. This fact makes translation of a policy from the fluid model to the CRW model relatively transparent: See the discussion of Case II for the routing model in Section 4, and the analysis of the MaxWeight policy in Section 6 and [9, 34]. This is not true when the cost is linear. In this case, $x^*_i = 0$ for all but one value of i when $w \ne 0$. If buffer i^* does not reside at station j then one must use some form of safety-stock to avoid starvation at station j.

In each of the examples that follow we have by assumption $i^* = 1$, so that all inventory lies at the first buffer in the relaxation. An approximate translation of this policy to the CRW model is obtained using the following switching curve,

$$s_\gamma(x_1) := \gamma \log(1 + x_1\gamma^{-1}), \qquad x_1 \ge 0. \tag{23}$$

When the total inventory at the bottleneck station j satisfies at time k,

$$\sum_{i:C_{ji}=1} Q_i(k) \le s_\gamma(Q_1(k))$$

then there is risk of starvation at station j, and the policy is designed to move customers from buffer 1 to this station. In the following three sections we describe in detail the construction of a policy, and appropriate values for the parameter γ.

To apply (22) we must compute the average effective cost for the relaxation defined in (19).

Proposition 2.1. *Consider the general one-dimensional queueing model (19) under the non-idling policy satisfying* $\rho_j = \lambda_j/\mu_j < 1$. *Let* $\sigma^2 := \rho m^2 + (1-\rho)m_L^2 < \infty$, *where* $m^2 = \mathsf{E}[(S_j(1) - L_j(1))^2]$ *and* $m_L^2 = \mathsf{E}[L(1)^2]$. *For a given constant* $\overline{c}_* > 0$, *the steady-state mean of* $\overline{c}_* \widehat{W}_j^*(k)$ *and the associated relative value function are expressed,*

$$\widehat{\eta}_* = \tfrac{1}{2}\overline{c}_* \frac{\sigma_j^2}{\mu_j - \lambda_j} \tag{24}$$

$$\hat{h}_*(w) = \tfrac{1}{2}\overline{c}_* \frac{w^2 + d_* w}{\mu_j - \lambda_j}, \qquad w \in \mathbb{R}_+, \tag{25}$$

where $d_* = \mu_j^{-1}(m^2 - m_L^2)$. *The function* $\hat{h}_*$ *solves a version of Poisson's equation for (19),*

$$\mathsf{E}[\hat{h}_*(\widehat{W}_j^*(k+1)) \mid \widehat{W}_j^*(k)] = \hat{h}_*(\widehat{W}_j^*(k)) - \overline{c}_* \widehat{W}_j^*(k) + \widehat{\eta}_*, \qquad k \ge 0.$$

Moreover, a similar identity holds for the unrelaxed process (18): For any allocation sequence $\boldsymbol{U}$ *satisfying (A2),*

$$\begin{aligned}\mathsf{E}[\hat{h}_*(W_j(k+1)) \mid Q(k)] = \hat{h}_*(W_j(k)) - \overline{c}_* W_j(k) + \widehat{\eta}_* \\ + \overline{c}_* \frac{\mu_j}{\mu_j - \lambda_j} W_j(k) I_j(k), \qquad k \ge 0.\end{aligned} \tag{26}$$

Proof. The representation of $\widehat{\eta}_*$ and $\hat{h}_*$ follow directly from Proposition 1.1 since (19) is a version of the simple queue.

The proof of (26) follows from the following computations,

$$\begin{aligned}\mathsf{E}[W_j(k+1) - W_j(k) \mid Q(k)] &= \lambda_j - \mu_j + \mu_j I_j(k) \\ \mathsf{E}[W_j^2(k+1) - W_j^2(k) \mid Q(k)] &= 2W_j(k)\mathsf{E}[W_j(k+1) - W_j(k) \mid Q(k)] \\ &\quad + \mathsf{E}[(W_j(k+1) - W_j(k))^2 \mid Q(k)] \\ &= 2W_j(k)(\lambda_j - \mu_j + \mu_j I_j(k)) + m^2 + I_j(k)(m_L^2 - m^2).\end{aligned}$$

The identity (26) thus follows from (25). □

We conclude this section with a central result providing criteria for fluid-scale and heavy-traffic asymptotic optimality, based on the existence of a Lyapunov function satisfying (11).

We consider a family of models parameterized by $\vartheta \in [0,1]$. It is assumed that there is a single bottleneck: for some fixed $\vartheta_0 \in (0,1)$, there is a unique integer $j \in \{1, \dots, \ell_r\}$ satisfying

$\rho_j = \rho_\bullet = \vartheta$ for each $\vartheta \in [\vartheta_0, 1]$. The dependency of $\boldsymbol{W}$ and $\boldsymbol{Q}$ on the parameter ϑ is suppressed to simplify notation.

The cost function is assumed to be linear throughout the remainder of this paper. It is also independent of ϑ. The effective cost is thus of the form $\overline{c}(w) = c(x^*(w)) = (c_{i^*}/\xi^j_{i^*})w$, $w \in \mathbb{R}_+$. By choice of indices we may assume without loss of generality that $i^* = 1$.

We list here the remaining assumptions on the policy and the parameterized model required in Theorem 2.2. These properties will be verified in the examples that follow for a fixed policy independent of ϑ.

(P1) For each $\vartheta \in [\vartheta_0, 1]$ the network satisfies assumptions (A1)–(A3), where the routing matrix R and the policy are independent of ϑ.

(P2) The random variables $\{A^\vartheta(k), S^\vartheta(k) : k \geq 1, \vartheta \in [\vartheta_0, 1]\}$ are defined on a common probability space, and are monotone in ϑ:

$$S^\vartheta(k) \downarrow S^1(k), \quad A^\vartheta(k) \uparrow A^1(k), \qquad \vartheta \uparrow 1, \text{ a.s., } k \geq 1.$$

(P3) The randomized stationary policy gives rise to a uniformly 0-irreducible Markov chain: for each $N \geq 1$ we can find $\varepsilon_N > 0$, $T_N \geq 1$ such that for all $\|x\| \leq N$, $\vartheta \in [\vartheta_0, 1]$,

$$\mathsf{P}_x\{Q(T_N) = 0\} \geq \varepsilon_N.$$

(P4) For each $\vartheta \in [\vartheta_0, 1)$ there exists a Lyapunov function $V = V^\vartheta$ satisfying (11). The Lyapunov functions and error constants satisfy the uniform bounds for $x \in \mathbb{N}^\ell$ and $\vartheta \in [\vartheta_0, 1)$,

$$\begin{aligned} J_*(x) \leq V(x) &\leq J_*(x) + (1-\vartheta)^{-1} K_{\mathrm{P4}}(1 + \|x\| \log(1 + \|x\|)), \\ \mathcal{E}(x) &= K_{\mathrm{P4}}\big(\log(1 + c(x)) + (1-\vartheta)^{-1}\big(1 + \langle \xi^j, x\rangle\big)^{-2}\big), \end{aligned}$$

where $K_{\mathrm{P4}} < \infty$ is independent of ϑ, and $J_* = J_*^\vartheta$ is the optimal fluid value function.

For each ϑ we denote by η the steady state cost under the given policy for the CRW model, and $\widehat{\eta}_*$ the optimal average cost for the one-dimensional relaxation. Proposition 2.1 implies the formula,

$$\widehat{\eta}_* = \tfrac{1}{2}\frac{1}{\mu_j}\frac{\sigma_j^2}{1-\vartheta}\frac{c_1}{\xi^j_1}, \qquad \vartheta_0 \leq \vartheta \leq 1, \tag{27}$$

where $\sigma_j^2 := \vartheta \mathsf{E}[(S_j^\vartheta(1) - L_j^\vartheta(1))^2] + (1-\vartheta)\mathsf{E}[(L_j^\vartheta(1))^2]$.

The bound (30) presented in Theorem 2.2 is a strong version of FSAO. We say that the family of networks is *heavy-traffic asymptotically optimal* (HTAO) *with logarithmic regret* when the bound (31) holds.

The proof of Theorem 2.2 is contained in the Appendix. We note that the starting point of the proof is the following consequence of the bound (11) and the Comparison Theorem of [31]: for any stopping time τ and any initial condition,

$$\mathsf{E}_x\Big[\sum_{k=0}^{\tau-1} c((Q(k)))\Big] \leq V(x) + \mathsf{E}_x\Big[\sum_{k=0}^{\tau-1} \widehat{\eta}_* + \mathcal{E}(Q(k))\Big]. \tag{28}$$

Part (i) is obtained on setting $\tau = nT$ for a given $T > 0$, $n \geq 1$, and applying bounds obtained in the Appendix. To prove Part (ii) we set $\tau = n$, divide both sides of the resulting bound by n, and let $n \to \infty$ to obtain,

$$\eta := \mathsf{E}_\pi[c(Q(k))] \leq \widehat{\eta}_* + \mathsf{E}_\pi[\mathcal{E}(Q(k))], \qquad k \geq 0. \tag{29}$$

The proof of (31) is based on uniform bounds on the steady-state mean of $\mathcal{E}$.

Theorem 2.2. *Suppose that (P1)–(P4) hold for some fixed policy, and some fixed linear cost function. Then, the following bounds hold for some $K_{2.2} < \infty$ that is independent of ϑ, and each $\vartheta \in [\vartheta_0, 1)$:*

(i) *The policy is FSAO, and the following explicit bound holds,*

$$0 \leq \limsup_{n\to\infty} \frac{n}{\log(n)} \Big(J_n(x,T) - J_*(x)\Big) \leq K_{2.2} \frac{\|x\|}{1-\rho_\bullet}, \qquad x \in \mathbb{R}^\ell_+,\ T \geq T_\circ(x), \tag{30}$$

where $T_\circ(x)$ is the draining time for the fluid model (20), of order $\|x\|(1-\rho_\bullet)^{-1}$.

(ii) *There is a unique invariant probability distribution π on $\mathbb{N}^\ell$, and the following bounds hold for the steady state cost:*

$$K_{2.2}(1-\rho_\bullet)^{-1} \leq \widehat{\eta}_* \leq \eta_* \leq \eta \leq \widehat{\eta}_* + K_{2.2}|\log(1-\rho_\bullet)|, \tag{31}$$

where η_* denotes the optimal steady-state cost for the CRW model (12). □

In the following three sections we illustrate and extend Theorem 2.2 in several examples.

3 Tandem queues

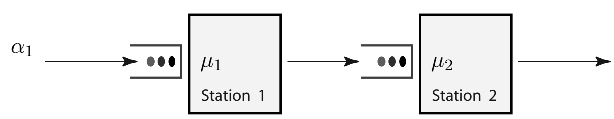


Figure 2: Two queues in tandem.

The model shown in Figure 2 was considered recently in [10] in their analysis of value functions for networks. We assume as in that paper that $c_1 < c_2$ and $\mu_1 > \mu_2$. For the fluid model there is a *pathwise* optimal policy that maintains $\zeta_1(t) = 0$ until buffer two is empty. Thereafter, the second buffer remains empty, but $\zeta_1(t) = \mu_2/\mu_1$ and $\zeta_2(t) = 1$ until the system is empty.

We focus on the second workload process since $\rho_1 < \rho_2$. This is given by the sum $w_2(t) = q_1(t) + q_2(t)$ in the fluid model, and satisfies the lower bound $\frac{d^+}{dt} w_2(t) \geq -(\mu_2 - \alpha_1)$ under any policy. The lower bound is achieved if, and only if $\zeta_2(t) = 1$. Its relaxation is defined by the controlled differential equation, $\frac{d^+}{dt}\widehat{w}_2(t) = -(\mu_2 - \alpha_1) + \iota(t)$, where $\iota(t) \geq 0$ for all $t \geq 0$, and the workload process $\widehat{\boldsymbol{w}}_2$ is non-negative. The effective cost is given by (21) which has the solution $\overline{c}(w) = c_1 w$ for $w \in \mathbb{R}_+$.

The value function for the fluid relaxation can be expressed as a function of workload $w = x_1 + x_2$,

$$\widehat{J}_*(w) = \tfrac{1}{2}c_1 \frac{w^2}{\mu_2 - \alpha_1}, \qquad w \in \mathbb{R}_+.$$

The fluid value function for $\boldsymbol{q}$ is given by the following perturbation:

$$J_*(x) = \widehat{J}_*(w_2(x)) + \tfrac{1}{2}(c_2 - c_1)\frac{x_2^2}{\mu_2}, \qquad x \in \mathbb{R}_+^2, \tag{32}$$

where $w_2(x) = x_1 + x_2$.

As for the stochastic network, we first consider a model satisfying (A1)–(A3) obtained from a model with Poisson arrivals and exponential service via uniformization. This is of the form (12) where the three-dimensional process $D(k) := (S_1(k), S_2(k), A_1(k))^{\mathrm{T}}$ is i.i.d. with marginal distribution given by,

$$\begin{aligned} \mathsf{P}\{D(k) = e^1\} &= \mu_1; \qquad \mathsf{P}\{D(k) = e^2\} = \mu_2; \\ \mathsf{P}\{D(k) = e^3\} &= \alpha_1, \end{aligned} \tag{33}$$

where e^j denotes the jth standard basis vector in $\mathbb{R}^3$. It is assumed that $\mu_1 + \mu_2 + \alpha_1 = 1$.

The workload process corresponding to the second station is defined as $W_2(k) = Q_1(k) + Q_2(k)$, $k \ge 0$. Its relaxation is a version of (19) with $j = 2$, and $L_2(k) = A_1(k)$. The idleness process $\hat{\boldsymbol{I}}_2$ is adapted to $\{(A_1(k), S_1(k), S_2(k)) : k \ge 1\}$, and is non-negative.

The one-dimensional relaxation has mean drift $\alpha_1 - \mu_2$, and second order parameters given by $m^2 = \mu_2 + \alpha_1$ and $m_A^2 = \alpha_1$. The steady-state mean of $\overline{c}(\widehat{W}_2^*(k))$ and the relative value function are obtained as an application of Proposition 2.1,

$$\widehat{\eta}_* := \lim_{k\to\infty} \mathsf{E}[\overline{c}(\widehat{W}_2^*(k))] = \frac{\alpha_1}{\mu_2 - \alpha_1} c_1, \tag{34}$$

$$\hat{h}_*(w) = \widehat{J}_*(w) + \tfrac{1}{2}c_1 \frac{w}{\mu_2 - \alpha_1} = \tfrac{1}{2}c_1 \frac{w^2 + w}{\mu_2 - \alpha_1}, \qquad w \in \mathbb{N}. \tag{35}$$

We now define a stationary policy for the CRW model using the switching curve $s_\gamma \colon \mathbb{R}_+ \to \mathbb{R}_+$ defined in (23). The policy at station 2 is non-idling, so that $U_2(k)1$ whenever the second buffer is non-empty. The server at station 1 works if $x_2 \le s_\gamma(x_1)$, which is expressed as the stationary policy,

$$U_1(k) = \mathbf{1}(Q_2(k) \le s_\gamma(Q_1(k)) \ \textit{and}\ Q_1(k) \ge 1), \qquad k \ge 0. \tag{36}$$

This is a minor modification of the policy proposed in [10] - the constant γ^{-1} within the logarithm in (23) is introduced to ensure a moderate slope at the origin. If $\gamma > 0$ is chosen sufficiently large then station 2 is rarely starved of work, and the policy ensures that the set $\mathsf{X} = \{x \in \mathbb{N}^2 : x_2 \le 2 + s_\gamma(x_1)\}$ is *absorbing* for the controlled process, so that if $Q(0) \in \mathsf{X}$, then $Q(k) \in \mathsf{X}$ for all $k \ge 0$. Consequently, the second queue is much smaller than the first whenever the system is congested.

Consider a family of networks as described in Section 2: The service rate $\mu_1 \in (\frac{1}{3}, \frac{1}{2})$ is fixed, we define $\vartheta_0 = \frac{3}{2}(1 - \mu_1)$, and for $\vartheta \in [\vartheta_0, 1)$ we define,

$$\mu_2^\vartheta = \mu_2^1 + (1 - \vartheta), \quad \alpha_1^\vartheta = \alpha_1^1 - (1 - \vartheta),$$

with $\mu_2^1 = \alpha_1^1 = \frac{1}{2}(1-\mu_1)$. When $\vartheta = \vartheta_0$ we have $\rho_1 = \rho_2 = \bar{\rho}_1 := \frac{1}{2}(1-\mu_1)/\mu_1 < 1$.

The proof of Proposition 3.1 is based on a Lyapunov function of the form (10), where J_* is defined in (32), and the function $b\colon \mathbb{N}^2 \to \mathbb{R}_+$ is defined as the correction term,

$$b(x) = c_1\mu_2\Big(\frac{1}{\mu_2-\alpha_1}\Big)\Big(\frac{1}{\mu_1-\mu_2}\Big)x_1 e^{-\beta x_2}, \qquad x \in \mathbb{N}^2\,, \tag{37}$$

with $\beta = \log(\mu_1/\mu_2)$. When $\vartheta = 1$ the constant β takes the maximum value

$$\bar{\beta} = \log(\mu_1/\mu_2^1) = \log(\mu_1/\alpha_1^1) = |\log(\bar{\rho}_1)|. \tag{38}$$

Proposition 3.1. *Consider the family of tandem queues controlled using the policy (36). It is assumed that γ is independent of ϑ, and satisfies,*

$$\gamma > 3/\bar{\beta},$$

where $\bar{\beta}$ is defined in (38). Then (P1)–(P4) hold, and hence the conclusions of Theorem 2.2 are satisfied with $\widehat{\eta}_*$ given in (34).

FSAO was obtained previously in [10] for this model using a policy of the same form, with a slightly different switching curve defined by $s(x_1) = \gamma\log(1+x_1)$ for $x_1 \ge 0$. The general form (23) is adopted here to ensure that the slope of s_γ at zero is independent of γ. The bound obtained on γ in [10, Theorem 2] is not easily compared to the bound assumed in Proposition 3.1, but both are unbounded as $\rho_1 \to 1$.

Proof of Proposition 3.1. The only part that requires proof is Property (P4). This requires consideration of the conditional expectation bound expressed in (11), which can be equivalently expressed,

$$PV\,(x) = V(x) - c(x) + \widehat{\eta}_* + \mathcal{E}(x), \qquad x \in \mathbb{N}^\ell, \tag{39}$$

where $V = J_* + b$ is defined in (10), and for any function h on the state space we define $Ph\,(x) := \sum_y P(x,y)h(y)$.

Consider first the transition kernel P applied to the fluid value function without the correction term: For $x \in \mathbb{N}^2$ we have,

$$\begin{aligned} PJ_*\,(x) &= \mathsf{E}[J_*(Q(k+1)) \mid Q(k) = x] \\ &= \tfrac{1}{2}\frac{c_1}{\mu_2-\alpha_1}\mathsf{E}[(w_2+\Delta_{W_2})^2] + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\mathsf{E}[(x_2+\Delta_{Q_2})^2] \end{aligned} \tag{40}$$

where $w_2 = x_1 + x_2$, $\Delta_{W_2} = W_2(k+1) - W_2(k)$, and $\Delta_{Q_2} = Q_2(k+1) - Q_2(k)$. Writing $U(k) = (u_1, u_2)^{\mathrm{T}} \in \{0,1\}^2$, the first and second moments are given by,

$$\begin{aligned} \mathsf{E}[\Delta_{W_2} \mid Q(k) = x] &= \alpha_1 - \mu_2 u_2 \qquad & \mathsf{E}[\Delta_{W_2}^2 \mid Q(k) = x] &= \alpha_1 + \mu_2 u_2 \\ \mathsf{E}[\Delta_{Q_2} \mid Q(k) = x] &= \mu_1 u_1 - \mu_2 u_2 \qquad & \mathsf{E}[\Delta_{Q_2}^2 \mid Q(k) = x] &= \mu_1 u_1 + \mu_2 u_2 \end{aligned}$$

We then obtain using (40) together with the identity $w_2 u_2 = (x_1 + x_2)u_2 = x_1 u_2$,

$$\begin{aligned} PJ_*\,(x) - J_*(x) &= \tfrac{1}{2}\frac{c_1}{\mu_2-\alpha_1}\Big(2w_2(\alpha_1-\mu_2u_2) + \alpha_1 + \mu_2u_2\Big) \\ &\quad + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\Big(2x_2(\mu_1u_1-\mu_2u_2) + \mu_1u_1 + \mu_2u_2\Big) \\ &= \tfrac{1}{2}\frac{c_1}{\mu_2-\alpha_1}\Big(2x_1\mu_2(1-u_2) + 2\alpha_1 + (\mu_2-\alpha_1) - \mu_2(1-u_2)\Big) \\ &\quad + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\Big(2x_2\mu_1u_1 + \mu_1u_1 + \mu_2u_2\Big) - c(x) \end{aligned}$$

Based on the formula (34) for $\widehat{\eta}_*$ we obtain,

$$
\begin{aligned}
PJ_*(x) &- [J_*(x) - c(x) + \widehat{\eta}_*] \\
&= \tfrac{1}{2}\frac{c_1}{\mu_2-\alpha_1}(2x_1-1)\mu_2(1-u_2) + \tfrac{1}{2}c_1 \\
&\qquad + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}(2x_2+1)\mu_1 u_1 + \tfrac{1}{2}(c_2-c_1) - \tfrac{1}{2}(c_2-c_1)(1-u_2) \\
&= \tfrac{1}{2}\frac{1}{\mu_2-\alpha_1}\Big(2c_1\mu_2 x_1 - c_1\mu_2 - (c_2-c_1)(\mu_2-\alpha_1)\Big)(1-u_2) \\
&\qquad + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\big(2x_2+1\big)\mu_1 u_1 + \tfrac{1}{2}c_1
\end{aligned}
\tag{41}
$$

We conclude that the drift inequality (11) holds using $V = J_*$ only for those states satisfying $x_2 \geq 1$. When $x_2 = 0$ then $u_2 = 0$, and the potentially large term $(\mu_2-\alpha_1)^{-1}c_1\mu_2 x_1$ is introduced. The function b is included in the definition (10) to counteract this.

Write $b_0(x) := e^{-\beta x_2} = (\mu_2/\mu_1)^{x_2}$ and $b_1(x) = x_1 b_0(x)$, $x \in \mathbb{N}^2$. Simple calculations reveal a negative drift whenever the second buffer is empty:

$$
\begin{aligned}
Pb_0(x) &= b_0(x) - (\mu_1-\mu_2); \\
Pb_1(x) &= b_1(x) - (\mu_2-\alpha_1) - (\mu_1-\mu_2)x_1, \qquad x_2 = 0.
\end{aligned}
\tag{42}
$$

The coefficient of b_1 in the function b is chosen so that $Pb(x) - b(x) = -(\mu_2-\alpha_1)^{-1}c_1\mu_2 x_1$ whenever $x_2 = 0$, $x_1 \neq 0$, thus canceling the positive contribution observed in (41) when $u_2 = 0$.

The function b_0 is *harmonic* near the lower boundary:

$$
\begin{aligned}
Pb_0(x) &= b_0(x) && 1 \leq x_2 \leq s_\gamma(x_1); \\
&= b_0(x) + \mu_2(e^\beta - 1)b_0(x) && x_2 > s_\gamma(x_1),
\end{aligned}
$$

and direct calculations provide similar expressions for the conditional mean of b_1,

$$
\begin{aligned}
Pb_1(x) &= b_1(x) - (\mu_2-\alpha_1)b_0(x) && 1 \leq x_2 \leq s_\gamma(x_1); \\
&= b_1(x) + \alpha_1 b_0(x) + \mu_2(e^\beta - 1)b_1(x) && x_2 > s_\gamma(x_1).
\end{aligned}
\tag{43}
$$

Moreover, we have the pair of bounds,

$$
b_0(x) \leq e^{-\beta\gamma\log(1+x_1/\gamma)} \leq (x_1/\gamma)^{-\beta\gamma}, \quad b_1(x) \leq x_1(x_1/\gamma)^{-\beta\gamma}, \qquad \text{when } x_2 > s_\gamma(x_1), \tag{44}
$$

and by assumption we have $\bar{\beta}\gamma > 3$, and hence $\beta\gamma \geq 3$ for all $\vartheta < 1$ sufficiently large.

Combining the expression (37) for b with (41) gives,

$$
\begin{aligned}
PV(x) - [V(x) - c(x) + \widehat{\eta}_*] \leq{}& \tfrac{1}{2}\frac{1}{\mu_2-\alpha_1}\big(2c_1\mu_2 x_2\big)(1-u_2) \\
&+ \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\big(2x_2+1\big)\mu_1 u_1 + \tfrac{1}{2}c_1 \\
&+ c_1\mu_2\Big(\frac{1}{\mu_2-\alpha_1}\Big)\Big(\frac{1}{\mu_1-\mu_2}\Big)\Big(Pb_1(x) - b_1(x)\Big).
\end{aligned}
$$

Substituting the expression for Pb_1 given in (43) and (42) then gives,

$$
\begin{aligned}
&PV(x) - [V(x) - c(x) + \widehat{\eta}_*] \\
&\qquad \le \tfrac{1}{2}\frac{c_2 - c_1}{\mu_2}\big(2x_2+1\big)\mu_1 u_1 + \tfrac{1}{2}c_1 \\
&\qquad\qquad + c_1\mu_2\Big(\frac{1}{\mu_2-\alpha_1}\Big)\Big(\frac{1}{\mu_1-\mu_2}\Big)\Big(\alpha_1 b_0(x) + \mu_1(e^{\beta}-1)b_1(x)\Big)(1-u_1)
\end{aligned}
$$

The coefficient of u_1 on the right hand side is $O(\log(x_1))$ since $x_2 \le s_\gamma(x_1)$ whenever $u_1 = 1$. To bound the coefficient of $(1-u_1)$ we note that $x_2 \ge s_\gamma(x_1)$ whenever $u_1 = 0$. From (44) it follows that $b_0(x) + b_1(x) = O\big((1+x_1+x_2)^{-2}\big)$ whenever $\beta\gamma \ge 3$, and this completes the verification of (P4). ☐

The same method of proof provides similar bounds for general service and arrival processes satisfying (A1). This is most easily seen when $\boldsymbol{A}$ and $\boldsymbol{S}$ are independent: Define the function b_0 as above, with β chosen so that $Pb_0 = b_0$ when $u_1 = u_2 = 1$. Hence β is the solution to the equation,

$$\Lambda(\beta) := \log(\mathsf{E}[\exp^{-\beta(S_1(k)-S_2(k))}]) = 0.$$

The function Λ is convex, with $\Lambda(0) = 0$ and $\Lambda'(0) = \mathsf{E}[-(S_1(k) - S_2(k))] < 0$, which implies that $\beta > 0$.

When $x_2 = 0$, $x_1 \ge 1$, we have the following analog of (42) when $\boldsymbol{A}$ and $\boldsymbol{S}$ are independent:

$$
\begin{aligned}
Pb_1(x) &= \mathsf{E}\big[(x_1 + A_1(1) - S_1(1))e^{-\beta(x_2+S_1(1))}\big] \\
&= b_1(x) - \mathsf{E}[1 - e^{-\beta S_1(1)}]x_1 + \mathsf{E}\big[(A_1(1) - S_1(1))e^{\beta S_1(1)}\big] \\
&= b_1(x) - \mu_1\big(1 - e^{-\beta}\big)x_1 - \big((1-\alpha_1)\mu_1 e^{\beta} - (1-\mu_1)\alpha_1\big).
\end{aligned}
$$

Hence a solution to (11) may be obtained as above with only superficial changes since $\beta > 0$.

4 Routing model

We now return to the routing model shown in Figure 1. To obtain a simplified model we set $\mu_3 = \infty$. On eliminating buffer 3 we arrive at the following two-dimensional stochastic network model,

$$
\begin{aligned}
Q(k+1) = Q(k) &- S_1(k+1)U_1(k)e^1 - S_2(k+1)U_2(k)e^2 \\
&+ A_r(k+1)U_1^r(k)e^1 + A_r(k+1)U_2^r(k)e^2, \qquad k \ge 0.
\end{aligned}
$$

It is assumed that (A1)–(A3) are satisfied, and that $S_1S_2 = 0$ a.s.. That is, at most one service is completed in each time slot.

The allocation process is subject to the *equality constraint* that $U_1^r(k) + U_2^r(k) = 1$ for $k \ge 0$. This arises from the constraint that the buffer at the router is always empty.

This model falls outside of the precise setting of Theorem 2.2. However, the conclusions can be carried over to this model based on the bound (28) once we verify Property (P4) for a parameterized family of models.

It is assumed throughout that $\mu_i < \alpha_r$ for each i so that it is necessary that the two stations cooperate to clear congestion. This is an example of *resource pooling* [18, 23]: the network load is not specified by a single station, but rather is given by $\rho_\bullet = \alpha_r/(\mu_1+\mu_2)$, and the workload is defined by $W(k) = Q_1(k)+Q_2(k)$ for $k \ge 0$. Its relaxation is defined for $k \ge 0$ by the recursion,

$$\widehat{W}(k+1) = \widehat{W}(k) + A_r(k+1) - (S_1(k+1) + S_2(k+1)) + (S_1(k+1) + S_2(k+1))\hat{I}(k).$$

The idleness process $\hat{\boldsymbol{I}}$ satisfies the assumptions imposed in the previous example. It is non-negative, and adapted to $\widehat{\boldsymbol{W}}$. The relaxation for the fluid model is defined analogously. It is expressed as the solution to the differential equation,

$$\frac{d^+}{dt}\widehat{w}(t) = -(\mu_1 + \mu_2 - \alpha_r) + \iota(t), \qquad t \ge 0.$$

We again assume that there is a linear cost function associated with the two buffers, and we take $c_1 \le c_2$ (without loss of generality). The optimal value function for the relaxation as a function of workload is given by

$$\widehat{J}_*(w) = \tfrac{1}{2}c_1\frac{w^2}{\mu_1 + \mu_2 - \alpha_r}, \qquad w \in \mathbb{R}_+,$$

and the value function for $\boldsymbol{q}$ under the optimal policy is

$$J_*(x) = \widehat{J}_*(x_1 + x_2) + \tfrac{1}{2}(c_2 - c_1)\frac{x_2^2}{\mu_2}, \qquad x \in \mathbb{R}^2_+.$$

Proposition 2.1 provides the following formulae for the optimal average-cost and the relative value function for the one-dimensional stochastic model,

$$\widehat{\eta}_* = \tfrac{1}{2}c_1\frac{\sigma^2}{\mu_1 + \mu_2 - \alpha_r} \tag{45}$$

$$\hat{h}_*(w) = \tfrac{1}{2}c_1\frac{w^2 + d_* w}{\mu_1 + \mu_2 - \alpha_r}, \qquad w \in \mathbb{R}_+, \tag{46}$$

where $\sigma^2 = \rho_2 m^2 + (1-\rho_2)m_A^2$ and $d_* = \mu_2^{-1}(m^2 - m_A^2)$, with $m^2 = \mathsf{E}[(S_1(1) + S_2(1) - A_r(1))^2]$ and $m_A^2 = \mathsf{E}[(A_r(1))^2]$.

For optimization we consider two special cases:

Case I If $c_1 < c_2$ then the optimal policy for the (unrelaxed) fluid model sends all arrivals to buffer 1, up until the first time that $q_2(t) = 0$. From this time up until the emptying time for the network, the policy maintains $q_2(t) = 0$, but sends material to this buffer at rate μ_2, so that $\zeta_1 + \zeta_2 = 1$ for all $0 \le t < w/(\mu_1 + \mu_2 - \alpha_r)$. This second requirement implies that the policy is time-optimal, so that

$$q(t; x) = 0, \qquad t \ge T^*(x) := \frac{x}{\mu_1 + \mu_2 - \alpha_r}.$$

In the stochastic model with $c_1 < c_2$ we consider the policy defined at the router by,

$$U_2^r(k) = \mathbf{1}(Q_2(k) \le s_\gamma(Q_1(k))), \qquad k \ge 0. \tag{47}$$

The two stations are non-idling, so that $U_i(k) = \mathbf{1}(Q_i(k) \ge 1)$, $i = 1, 2$.

Case II When the cost parameters are equal then *any* time-optimal policy is optimal for the fluid model, and we have $J_*(x) = \widehat{J}_*(x_1 + x_2)$ for all x. In particular, the policy defined by the linear switching curve $s(x_1) = \varpi x_1$, $x_1 \ge 0$ is optimal, where the constant $\varpi > 0$ is fixed, but arbitrary. We see in Proposition 4.1 that the insensitivity found in the fluid model is reflected in the CRW model: when controlled using the linear switching curve with $\varpi \in (0, \infty)$, there exists $K_{4.1} < \infty$ such that,

$$\widehat{\eta}_* \le \eta \le \widehat{\eta}_* + K_{4.1}, \qquad 0 < \vartheta < 1. \tag{48}$$

In either case, the Lyapunov function considered in this model is defined as (10) with correction term given by,

$$b(x) = \frac{c_1}{\mu_1 + \mu_2 - \alpha_r} \Big(\big(\mu_1 x_2 + m_A^2\big) k_1 e^{-\beta_1 x_1} + \big(\mu_2 x_1 + m_A^2\big) k_2 e^{-\beta_2 x_2} \Big) \tag{49}$$

where $\{\beta_j > 0\}$ solve,

$$\mathsf{E}[\exp(\beta_j (S_j(k) - A_r(k)))] = 1, \qquad j = 1, 2, \tag{50}$$

and the constants are expressed $k_i = (1 - \mathsf{E}[e^{-\beta_i A_r(k)}])^{-1}$, $i = 1, 2$.

To formulate an analog of Proposition 3.1 we construct a one-dimensional family of models as follows: Suppose that $(\boldsymbol{A}_r^1, \boldsymbol{S}_1, \boldsymbol{S}_2)$ satisfies (A1), with $\mathsf{E}[A_r^1(k)] = \mu_1 + \mu_2$. For each $0 \le \vartheta \le 1$, define the thinning of this arrival stream by $A_r^\vartheta(k) = T(k) A_r^1(k)$, where $\boldsymbol{T}$ is an i.i.d. Bernoulli process that is indpendent of $(\boldsymbol{A}_r^1, \boldsymbol{S}_1, \boldsymbol{S}_2)$ with $\mathsf{P}\{T(k) = 1\} = \vartheta$. The system load is given by $\rho_\bullet^\vartheta = \vartheta$ for each $\vartheta \in [0, 1]$. We restrict to $\vartheta \ge \vartheta_0 := \mu_1/(\mu_1 + \mu_2)$ to maintain the restriction that $\mu_i \le \alpha_r$ for each i.

Note that here we must explicitly assume that $m^2 > 0$ when $\vartheta = 1$ to ensure that $\widehat{\eta}_*$ is unbounded as $\vartheta \uparrow 1$. Positivity of m^2 follows from (A1) under the assumptions of Theorem 2.2.

Proposition 4.1. *Suppose that $\gamma > 3/\bar{\beta}_2$, where $\bar{\beta}_2$ solves (50) with $j = 2$ and $A_r(k) = A_r^1(k)$. Suppose moreover that $\mathsf{E}[(S_1(1) + S_2(1) - A_r^1(1))^2] > 0$. Then, Properties (P2)–(P4) hold, and the following consequences hold for the controlled network:*

(i) *The conclusions of Theorem 2.2 hold, where $\widehat{\eta}_*$ is given in (45).*

(ii) *If $c_1 = c_2$ and the linear switching curve is used, then the bound (48) holds for some $K_{4.1} < \infty$.*

Proof. To establish part (i) it is again sufficient to establish (P4). The proof is very similar to the proof of Proposition 3.1.

First consider the transition kernel applied to the fluid value function: For $x \in \mathbb{N}^2$ we have,

$$\begin{aligned} PJ_*(x) &= \mathsf{E}[J_*(Q(k+1)) \mid Q(k) = x] \\ &= \tfrac{1}{2} \frac{c_1}{\mu_1 + \mu_2 - \alpha_r} \mathsf{E}[(w_2 + \Delta_{W_2})^2] + \tfrac{1}{2} \frac{c_2 - c_1}{\mu_2} \mathsf{E}[(x_2 + \Delta_{Q_2})^2] \end{aligned} \tag{51}$$

where the increments $\Delta_{W_2} = W_2(k+1) - W_2(k)$, $\Delta_{Q_2} = Q_2(k+1) - Q_2(k)$ satisfy,

$$\begin{aligned} \mathsf{E}[\Delta_{W_2}] &= \alpha_r - \mu_1 u_1 - \mu_2 u_2 & \mathsf{E}[\Delta_{W_2}^2] &\le m^2 + (1 - u_1 u_2)(\overline{m}^2 - m^2) \\ \mathsf{E}[\Delta_{Q_2}] &= \alpha_r u_2^r - \mu_2 u_2 & \mathsf{E}[\Delta_{Q_2}^2] &\le \overline{m}^2, \end{aligned}$$

where $\overline{m}^2 := \max_u \mathsf{E}[(u_1 S_1(1) + u_2 S_2(1) - A_r(1))^2]$, and each expectation is conditional on $Q(k) = x$ and $U(k) = u$. Combining these expressions with (51) we obtain the bound,

$$PJ_*(x) \le J_*(x) + \tfrac{1}{2}\frac{c_1}{\mu_1+\mu_2-\alpha_r}\Big(2w_2(\alpha_r - \mu_1 u_1 - \mu_2 u_2) + m^2 + (1-u_1u_2)(\overline{m}^2 - m^2)\Big)$$
$$+ \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\Big(2x_2(\alpha_r u_2^r - \mu_2 u_2) + \overline{m}^2\Big).$$

To extract $\widehat{\eta}_*$ from the right hand side note that $m^2 = \sigma^2 + (1-\rho_2)(m^2 - m_A^2)$. Hence from (45),

$$\begin{aligned} &PJ_*(x) - [J_*(x) - c(x) + \widehat{\eta}_*] \\ &\quad\le \tfrac{1}{2}\frac{c_1}{\mu_1+\mu_2-\alpha_r}\Big(2w_2(\mu_1(1-u_1)+\mu_2(1-u_2)) + (1-\rho_2)m^2 + (1-u_1u_2)\overline{m}^2\Big) \\ &\qquad + \tfrac{1}{2}\frac{c_2-c_1}{\mu_2}\Big(2x_2(\alpha_r u_2^r + \mu_2(1-u_2)) + \overline{m}^2\Big) \\ &\quad\le \tfrac{1}{2}\frac{c_1}{\mu_1+\mu_2-\alpha_r}\big(2\mu_1 x_2 + \overline{m}^2\big)(1-u_1) \\ &\qquad + \tfrac{1}{2}\frac{c_1}{\mu_1+\mu_2-\alpha_r}\big(2\mu_2 x_1 + \overline{m}^2\big)(1-u_2) + \frac{c_2-c_1}{\mu_2}x_2\alpha_r u_2^r + K, \end{aligned} \tag{52}$$

where $K = \frac{1}{2}\big((c_2-c_1)\mu_2^{-1} + c_1(\mu_1+\mu_2)^{-1}\big)\overline{m}^2$. Define for $i = 1, 2$,

$$b_{0i}(x) = e^{-\beta_i x_i}, \quad b_{1i}(x) = x_{\bar{\imath}} e^{-\beta_i x_i}, \qquad x \in \mathbb{N}^2,$$

where $\bar{\imath} := 1 + |2-i|$ is the complementary index. Exactly as in the proof of Proposition 3.1 we find that $Pb_{ji} - b_{ji}$ is non-positive whenever $u_i^r = 1$. In particular,

$$\begin{aligned} &Pb_{0i}(x) = b_{0i}(x) - \big(1 - \mathsf{E}[e^{-\beta_i A_r(k)}]\big) \\ \textit{and}\quad &Pb_{1i}(x) = \mathsf{E}[e^{-\beta_i A_r(k)}]b_{1i}(x) - u_{\bar{\imath}}\mathsf{E}[S_{\bar{\imath}}(k)e^{-\beta_i A_r(k)}] \\ &\qquad\quad \le b_{1i}(x) - \big(1 - \mathsf{E}[e^{-\beta_i A_r(k)}]\big)x_{\bar{\imath}} \qquad \textit{whenever } x_i = 0 \textit{ and } u_i^r = 1. \end{aligned} \tag{53}$$

Combining the definition of b given in (49) with these bounds and (52), we obtain a bound of the form,

$$PV_*(x) - [V_*(x) - c(x) + \widehat{\eta}_*]$$
$$\le \frac{c_2-c_1}{\mu_2}x_2\alpha_r u_2^r + K + \mathcal{E}_0(x), \qquad x \in \mathbb{N}^2,$$

where the error term $\mathcal{E}_0$ appears due to the positive value of $Pb_{ji} - b_{ji}$ whenever $u_i^r = 0$ for $i = 1$ or $i = 2$. Exactly as in the proof of Proposition 3.1 it may be shown that $Pb_{ji} - b_{ji}$ is of order $(1 + x_1 + x_2)^{-2}$ when $u_i^r = 0$, and this completes the verification of (P4).

To see why the bound improves when $c_1 = c_2$ we strengthen the bound on the error term in (P4): it can be shown that for some constant $K_1 < \infty$ independent of ϑ,

$$\mathcal{E}(x) \le K_1(1-\vartheta)^{-1}\Big(x_1 e^{-\beta_2 x_2}\mathbf{1}(x_2 \ge \varpi x_1) + x_2 e^{-\beta_1 x_1}\mathbf{1}(x_2 \le \varpi x_1)\Big), \qquad x \in \mathbb{R}_+^2.$$

It follows that for a possibly larger constant, and some $\beta > 0$,

$$\mathcal{E}(x) \le K_1(1-\vartheta)^{-1}e^{-\beta w_2}, \qquad x \in \mathbb{R}^2_+. \tag{54}$$

We can define β explicitly by,

$$\beta = \min_{\vartheta\in[\vartheta_0,1]} \tfrac{1}{2}\min\{\beta_2^\vartheta, \varpi\beta_2^\vartheta, \beta_1^\vartheta, \varpi^{-1}\beta_1^\vartheta\} = \tfrac{1}{2}\min\{\beta_2^{\vartheta_0}, \varpi\beta_2^{\vartheta_0}, \beta_1^{\vartheta_0}, \varpi^{-1}\beta_1^{\vartheta_0}\}.$$

This is strictly positive since μ_1 and μ_2 are strictly positive.

Exploiting minimality of $\widehat{\boldsymbol{W}}^*$ gives the following bound on the steady state mean of the exponential,

$$\mathsf{E}_\pi[e^{-\beta W(k)}] \le \mathsf{E}_\pi[e^{-\beta \widehat{W}^*(k)}], \qquad k \ge 0.$$

The right hand side can be bounded using Lemma A.1 since $\widehat{\boldsymbol{W}}^*$ is a version of the simple queue:

$$\mathsf{E}_\pi[e^{-\beta \widehat{W}^*(k)}] \le (\mu_1+\mu_2-\vartheta\alpha_r^1)\frac{1}{\mathsf{E}[(A_r^\vartheta(1)-S_1(1)-S_2(1))_+]}\sum_{k=0}^\infty e^{-\beta k}.$$

The denominator is non-zero since $m^2 := \mathsf{E}[(A_r^\vartheta(1)-S_1(1)-S_2(1))^2] > 0$ for $\vartheta \in [\vartheta^0, 1]$. This together with (54) and the identity $\mu_1+\mu_2 = \alpha_r^1$ establishes the bound in (ii) with,

$$K_{4.1} = K_1\frac{\mu_1+\mu_2}{\mathsf{E}[(A_r^{\vartheta_0}(1)-S_1(1)-S_2(1))_+]}\frac{1}{1-e^{-\beta}}.$$

□

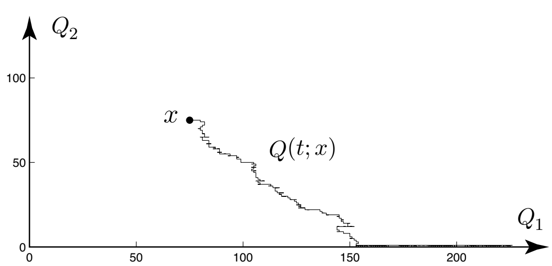
We conclude with numerical results to illustrate the need for safety-stocks in this model, and also to illustrate the conclusions of Proposition 4.1 under the proposed policy.

Let $n \ge 1$ denote a 'burstiness parameter', and with $D(k) := (S_1(k), S_2(k), A_r(k))^{\mathrm{T}}$ i.i.d., suppose that

$$\begin{aligned} &\mathsf{P}\{D(k)=e^1\} = \mu_1; \qquad \mathsf{P}\{D(k)=e^2\} = \mu_2;\\ &\mathsf{P}\{D(k)=ne^3\} = \alpha_r - \mathsf{P}\{D(k)=0\} = \alpha_r/n. \end{aligned} \tag{55}$$

It is assumed that $\mu_1+\mu_2+\alpha_r = 1$.

In the calculations that follow we take $\mu_1=\mu_2$, $\rho_\bullet = 0.9$, and the one-step cost is assumed linear with $c_1 = 2$ and $c_2 = 3$.



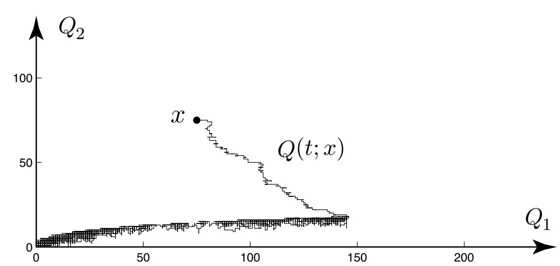


Figure 3: A naive translation of the optimal policy for the fluid model may not be stabilizing in a stochastic model. In the simulation shown at left, the queue length process $\boldsymbol{Q}$ explodes along the Q_1-axis. The simulation at right shows the policy described in Case I with $\gamma = 5$.

The nonlinear switching curve (23) is appropriate since $c_1 < c_2$. The formula determining β_2 in this model is expressed,

$$1 - (\mu_2 + \alpha_r n^{-1}) + \alpha_r n^{-1} e^{-n\beta_2} + \mu_2 e^{\beta_2} = 1, \tag{56}$$

which may be explicitly solved for $n = 1$, giving $\beta_2 = \log(\alpha_r/\mu_2)$.

For large n we can approximate the solution to (56) by $\beta_2 = an^{-1}$, where $a > 0$ is the solution to

$$\alpha_r n^{-1}(e^{-a} - 1) + \mu_2(e^{n^{-1}a} - 1) = 0.$$

Using the approximation $e^{n^{-1}a} \approx 1 + n^{-1}a$ and the formula $\rho_\bullet = \alpha_r/(\mu_1 + \mu_2) = \frac{1}{2}\alpha_r/\mu_2$ then gives,

$$\tfrac{1}{2}a - \rho_\bullet(1 - e^{-a}) \approx 0.$$

For $\rho_\bullet \sim 1$ we obtain $a \approx 1.5$, and the switching curve parameter is thus bounded from below by $\gamma \geq 3/\beta_2 \approx 2n$.

We illustrate the conclusions of Proposition 4.1 through numerical results previously presented in [28, Section 3.2]. This routing model was considered using the specific values $n = 1$, $\alpha_r = 9/19$, and $\mu_1 = \mu_2 = 5/19$. The foregoing calculations give $3\beta_2^{-1} = 3/\log(9/5) \approx 5.1$ for the lower bound on γ to obtain fluid-scale and heavy-traffic asymptotic optimality.

The grey region shown in Figure 1 indicates the optimal policy for the model with these parameters, obtained using value iteration. The switching curve s_γ with $\gamma = 5$ is also shown in this figure. It is a remarkably accurate approximation to the optimal policy. Numerical results reported in [38] also show a roughly logarithmic offset between policies for the fluid and stochastic models. A simulation of a sample path of $\boldsymbol{Q}$ under this policy is shown at right in Figure 3.

Consider for comparison the priority policy in which the router sends all customers to buffer one whenever buffer two is non-empty,

$$U_1^r(k) = \mathbf{1}(Q_2(k) \geq 1); \quad U_2^r(k) = 1 - U_1^r(k). \tag{57}$$

This is similar to the policy considered in Case I with $\gamma = 0$. As illustrated in the simulation at left in Figure 3, this policy does not perform well at all! The problem is, when $Q_2(k) = 0$ there is some delay before new work can be routed to the second queue. As can be seen in Figure 3, this delay will result in instability when the load is above some critical value.

5 The Dai-Wang network

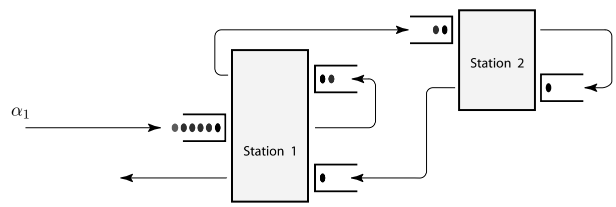


Figure 4: The five-buffer model of Dai and Wang.

The model shown in Figure 4 was introduced in [8] to show that a central limit scaling may not lead to a useful limit as the system load approaches unity. We consider this model here to show how to construct a policy for a network in which the majority of inventory is stored several steps away from the bottleneck. Unfortunately, there is no space for a detailed analysis of this model. Instead we describe briefly the construction of a policy and Lyapunov function of the form (10), and use simulation to illustrate the performance of the proposed policy.

The CRW model is considered under Assumptions (A1)–(A3). The two workload processes are defined by,

$$\begin{aligned} W_1(k) &= 3Q_1(k) + 2Q_2(k) + Q_3(k) + Q_4(k) + Q_5(k) \\ W_2(k) &= 2\big(Q_1(k) + Q_2(k) + Q_3(k)\big) + Q_4(k), \qquad k \geq 0. \end{aligned} \tag{58}$$

It is assumed that the second station is a bottleneck, so that $\rho_1 < \rho_2 = \rho_\bullet$, with $\rho_1 = 3\alpha_1/\mu_1$ and $\rho_2 = 2\alpha_1/\mu_2$.

Given any linear cost function on buffer levels, the effective cost based on the second workload process is given by $\bar{c}(w) = \bar{c}_* w$, with $\bar{c}_* = \min\{\frac{1}{2}c_1, \frac{1}{2}c_2, \frac{1}{2}c_3, c_4\}$. We assume that the minimum is achieved uniquely by the first term, so that $c_1 < \min\{c_2, c_3, 2c_4\}$, and $\bar{c}_* = \frac{1}{2}c_1$.

To mimic the optimal policy for the relaxation we wish to keep most inventory in buffer one. However, it is also necessary to feed station 2 when starvation is imminent. As in the previous examples, the switching curve s_γ is defined in (23), but in this example we define *two* regions in which starvation avoidance is prioritized:

$$\mathcal{P}_1 = \{x \in \mathbb{N}^5 : x_3 + x_4 \leq s_\gamma(x_1),\ x_2 \neq 0\}, \quad \mathcal{P}_2 = \{x \in \mathbb{N}^5 : x_2 + x_3 + x_4 \leq s_\gamma(x_1)\}. \tag{59}$$

The policy defined below is designed to move inventory from station 1 to station 2 when $Q(k) \in \mathcal{P}_1$, and from buffer 1 to buffer 2 when $Q(k) \in \mathcal{P}_2$. These goals are captured in the following two drift conditions: for some $\varepsilon_1, \varepsilon_2 > 0$, whenever $x_1 \neq 0$,

$$\begin{aligned} \mathsf{E}[Q_3(k+1) + Q_4(k+1) \mid Q(k) = x] &\geq x_3 + x_4 + \varepsilon_1, \qquad x \in \mathcal{P}_1, \\ \mathsf{E}[Q_2(k+1) \mid Q(k) = x] &\geq x_2 + \varepsilon_2, \qquad x \in \mathcal{P}_2. \end{aligned} \tag{60}$$

The following randomized policy is designed so that the bounds (60) hold with $\varepsilon_1 = \frac{2}{5}\mu_1 - \frac{1}{2}\mu_2$, and $\varepsilon_2 = \frac{1}{5}\mu_1$. We denote for each $i = 1, \ldots, 5$ and $x \in \mathbb{N}^5$,

$$u_i := \mathsf{P}\{U_i(k) = 1 \mid X(k) = x\}, \qquad k \geq 0.$$

(i) If $x \in \mathcal{P}_1$ then $u_2 = 2/5$. Otherwise, $u_2 = 0$.

(ii) If $x \in \mathcal{P}_2$ and $x_1 \geq 1$ then $u_1 = 3/5$. Otherwise, $u_1 = 0$.

(iii) If $x_5 \neq 0$ then $u_5 = 1 - u_1 - u_2$.

(iv) At station 2 the policy is defined as the non-idling, priority policy. Priority is given to buffer 3 if the cost parameters satisfy $c_3 > 2c_4$, and to buffer 4 otherwise.

Analysis of this policy can be performed as in the previous two examples by exploiting two facts: (i) the value function J_* is C^1, and (ii) $J_*(x) - \widehat{J}_*(x)$ is bounded as $\rho_2 \uparrow 1$ for each $x \in \mathbb{R}^5_+$ provided ρ_1 remains bounded away from unity. To verify (P4) we can again choose V of the form (11), where the function $b\colon \mathbb{R}^5_+ \to \mathbb{R}_+$ consists of two parts,

$$b(x) = k_1(x_1 + x_2)e^{-\beta_1(2x_3+x_4)} + k_2 x_1 e^{-\beta_2(x_1+2x_3+x_4)}.$$

The first term is introduced to provide a negative drift of order $(x_1+x_2)(1-\rho_2)^{-1}$ when station 2 is empty, and the second provides a negative drift when *buffer* 2 is empty. This leads to the following specifications: The exponents solve,

$$\begin{aligned} 1 &= \tfrac{2}{5}\mathsf{E}\big[e^{-\beta_1(-S_2(k)+2S_1(k))}\big] + \tfrac{3}{5}\mathsf{E}\big[e^{\beta_1 S_2(k)}\big], \\ 1 &= \mathsf{E}\big[e^{-\beta_2(S_1(k)-S_2(k))}\big], \end{aligned} \tag{61}$$

and the constants are of the form $k_i = (1-\rho_2)^{-1}k_i^0$. A bound of the form bound (P4) can be verified for a family of models satisfying these conditions. For HTAO we require $\beta_i^\vartheta \gamma \geq 3$ for $i = 1, 2$ and all $\vartheta \in [\vartheta_0, 1]$; the constants $\{k_i^0\}$ are independent of ϑ but must be sufficiently large; and we must have $\rho_1^\vartheta < 1$ when $\vartheta = 1$.

In simulations of this network under this policy the cost function was taken as $c(x) = x_1 + 2(x_2 + x_3 + x_4 + x_5)$, $x \in \mathbb{R}^\ell_+$. A family of CRW models was constructed exactly as in the routing model, where the common distribution of $D(k) := (S_1(k), S_2(k), A_1(k))^{\mathrm{T}}$ was specified by (55) for a given $n \geq 1$. The load at the first station was fixed at $\rho_1 = 0.8$, and ρ_2 was restricted to the range $[0.9, 1]$. The parameters $\{\beta_i\}$ defined in (61) are independent of n in this model.

Shown in Figure 5 are six plots obtained using $\rho_1 = 0.8$, $\rho_2 = 0.9$, $n = 1, 3, 5, 7, 9, 11$, and $\gamma = 1, 3, 5, 7, \dots, 19$. For each n, identical sample paths of the service and arrival processes were held fixed in the experiments using these ten values of γ, and the queue was initially empty, $Q(0) = 0$. The vertical axis shows the average of $c(Q(k))$ for $k \leq T = 10^5$. The optimal value of γ is very insensitive to the parameter n: In all but the first instance it lies between 3 and 5.

A second set of simulations were obtained with ρ_2 increased, and ρ_1 held fixed at 0.8. Shown in Figure 6 are results for $\rho_2 = 0.95$ and $\rho_2 = 1$. The arrival and service processes were again defined using (55), with burstiness parameter fixed at $n = 7$. The vertical axis again shows the average of $c(Q(k))$, but the time horizon was increased to $k \leq T = 10^6$ to account for the greater system load. Note that the shape of the plot is virtually unchanged as the load varies between the three values $\rho_2 = 0.9$, 0.95, 1.

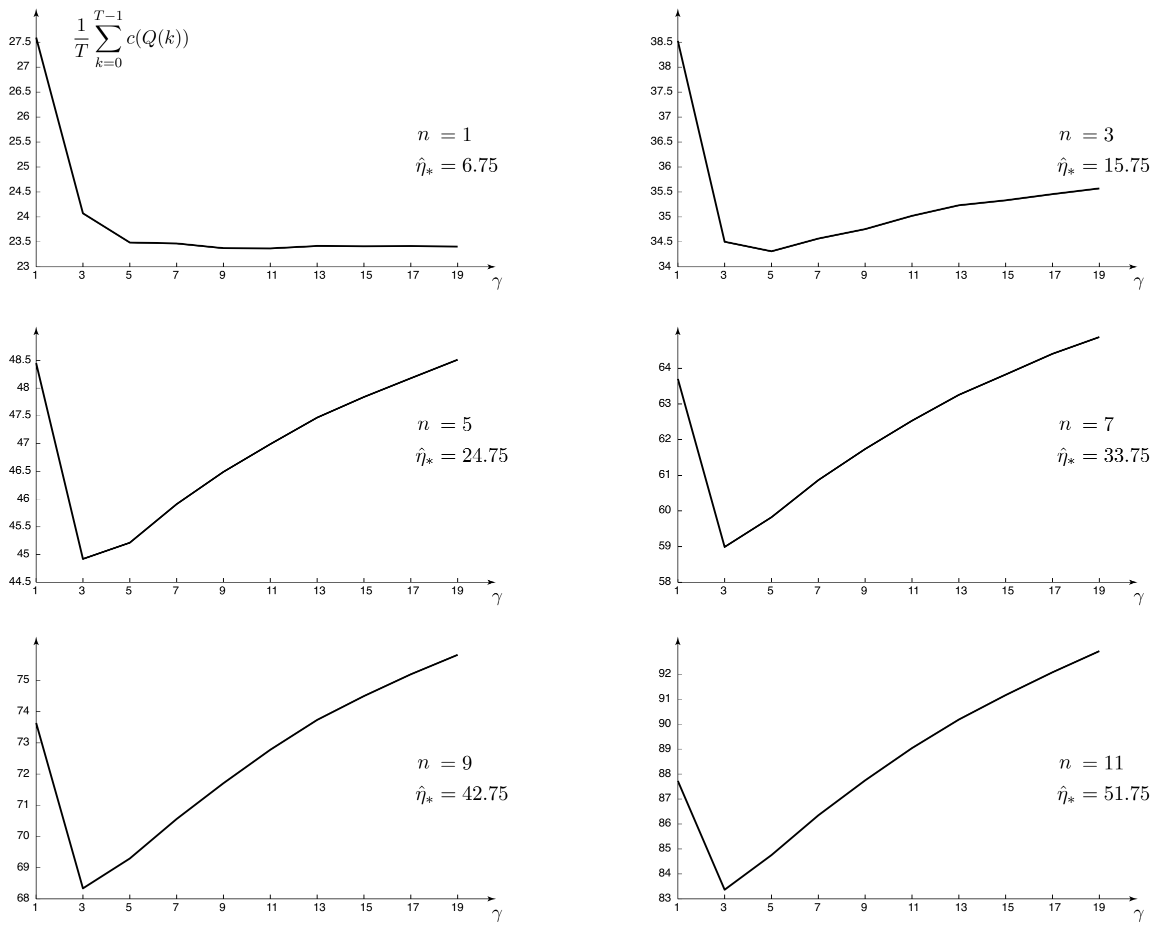


Figure 5: Average cost for the model of Dai and Wang with using $\rho_1 = 0.8$ and $\rho_2 = 0.9$.

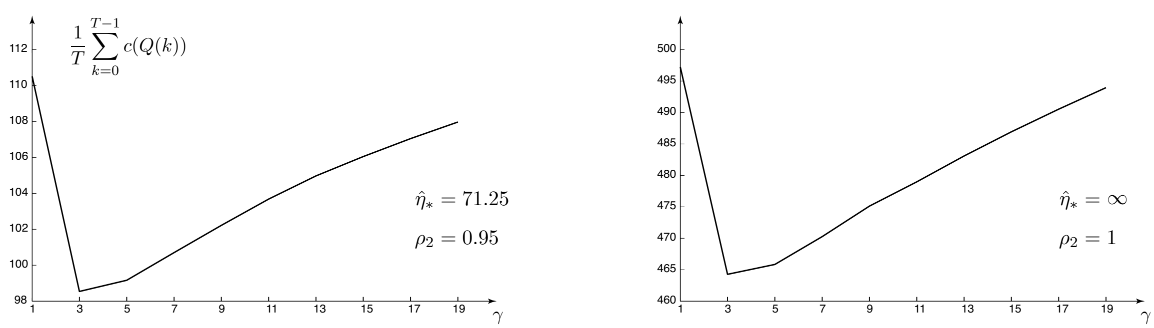


Figure 6: In these experiments the burstiness-parameter was fixed at $n = 7$, and the time horizon was taken to be $T = 10^6$. The load is $\rho_2 = 0.95$ in the plot shown at left, and $\rho_2 = 1$ in the plot shown at right. The policy is fixed throughout. It is precisely the same as used in the simulations shown in Figure 5.

6 Generalizations

Up to now it appears that the construction of an asymptotically optimal policy is a specialized art-form. In this section we introduce a general approach to obtain a stabilizing and FSAO policy based on the fluid value function. Although a completely general approach to HTAO is still lacking, some of the techniques obtained here may be useful in subsequent research.

On considering general multi-dimensional networks we exploit the fact that J_* is typically continuously differentiable (C^1) [30]. We shall take this for granted throughout this section.

Suppose that $V\colon \mathbb{R}^\ell_+ \to \mathbb{R}_+$ is any C^1 function that vanishes only at the origin. To define the myopic policy with respect to V, recall that $\mathsf{U}_0(x)$ is defined in (15) as the set of feasible allocation values for the CRW model when $Q(k) = x$. For the fluid model we denote by $\mathsf{U} = \{u \in \mathbb{R}^\ell_+ : Cu \le 1\}$ the set of feasible allocation rates, and the set of feasible allocations rates when the state takes the value $x \in \mathbb{R}^\ell_+$ is given by,

$$\mathsf{U}(x) = \{u \in \mathsf{U} : v_i := (Bu + \alpha)_i \ge 0 \text{ when } x_i = 0\}\,. \tag{62}$$

The myopic policy is defined for the fluid and stochastic models by the respective feedback laws,

$$f_V(x) = \mathop{\arg\min}_{u \in \mathsf{U}(x)} \langle \nabla V\,(x), Bu + \alpha\rangle, \tag{63}$$

$$f^0_V(x) = \mathop{\arg\min}_{u \in \mathsf{U}_0(x)} \mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x, U(k) = u]. \tag{64}$$

The myopic policy (63) is stabilizing for the *fluid model* under mild assumptions on the function V (see [6] and [4, Thm. 12.5] for linear functions, and [28, Proposition 11] for a smooth norm on $\mathbb{R}^\ell$.) The idea is that the function V serves as a Lyapunov function for the fluid model under the policy f_V. This construction fails for the stochastic model: as discussed in [28], the inclusion $\mathsf{U}_0(x) \subset \mathsf{U}(x)$ is strict, and hence the the mean increment $\mathsf{E}^{f^0_V}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x]$ may be positive for x on the boundaries of the state space.

Consider for example the quadratic function $V(x) = \frac{1}{2}x^{\mathrm{T}}Dx$ with $D > 0$ diagonal. Then $V^{\frac{1}{2}}$ is a norm on $\mathbb{R}^\ell$, so the policy f_V is stabilizing for the fluid model by [28, Proposition 11] (the myopic policies defined with respect to V or V^p are identical when $p > 0$.) The resulting feedback law f_V is a special case of the maximum pressure policy of [7], which is a generalization of the MaxWeight and related policies considered in [36, 11, 34, 25].

The main results of [36, 7] imply that f_V is a stabilizing policy for a stochastic network under very general assumptions. The analysis is similar to the aforementioned theory of the myopic policy for fluid models. The function V satisfies a drift condition for the stochastic model because it is convex and monotone, and also satisfies the following derivative condition,

$$\frac{\partial}{\partial x_i} V(x) = 0, \qquad \text{whenever } x_i = 0,\ i = 1, \dots, \ell. \tag{65}$$

This allows us to consider only $x \in \mathsf{U}_0(x)$ in the minimization (63) so that f_V is a feasible policy for the CRW model:

Proposition 6.1. *Suppose that the stochastic network satisfies (A1)–(A3), and that $V\colon \mathbb{R}^\ell \to \mathbb{R}_+$ is any C^1, monotone function satisfying the derivative condition (65). Then, the minimization (63) is equivalent to the restricted minimization over the set $\mathsf{U}_0(x)$.*

Proof. If $x_i = 0$ then $(Bu + \alpha)_i \geq 0$ for any feasible $u \in \mathsf{U}(x)$. Moreover, it follows from (65) that the inner product in (63) can be expressed,

$$\langle \nabla V\,(x), Bu + \alpha \rangle = \sum_{j \neq i} \Big(\frac{\partial}{\partial x_j} V(x) \Big) \Big(\alpha_j - \mu_j u_j + \sum_k \mu_k u_k R_{kj} \Big)$$

Monotonicity of V means that $\nabla V(x) \in \mathbb{R}_+$ for any $x \in \mathbb{R}_+^\ell$. It follows that the coefficients of u_i on the right hand side all are all non-negative, which implies that without loss of generality we may assume that $u_i = 0$ in an optimal solution:

$$f_V(x) = \arg\min\{\langle \nabla V\,(x), Bu \rangle \;:\; u \in \mathsf{U} \text{ and } u_i = 0 \text{ if } x_i = 0\}.$$

This completes the proof since the extreme points of this linear program are precisely $\mathsf{U}_0(x)$. ☐

We now introduce a function V based on the fluid model and a state-transformation that satisfies the assumptions of Proposition 6.1. Fix a constant $\gamma > 0$, and define for $x \in \mathbb{R}_+^\ell$ the vector $\tilde{x} \in \mathbb{R}_+^\ell$ via,

$$\tilde{x}_i = x_i + \gamma(e^{-x_i/\gamma} - 1)\,, \qquad i = 1, \dots, \ell.$$

In the remainder of this section we consider the myopic policy with respect to $V(x) = J_*(\tilde{x})$. The resulting policy is very similar to those considered in the previous three examples. In particular, it is straightforward to show that for the tandem queue or routing model, under the assumptions imposed above the feedback law f_V^0 is approximated by a logarithmic switching curve for large x_1.

Lemma 6.2. *Suppose that the fluid value function J_* is C^1. Then,*

(i) *The function V is convex and monotone.*

(ii) *The derivative condition (65) holds.*

(iii) $\lim_{n\to\infty} n^{-2} V(nx) = J_*(x)$ *for* $x \in \mathbb{R}_+^\ell$.

Proof. Part (i) follows from the fact that J_* has these properties (see [28, Proposition 6].)

The partial derivatives of V are given by,

$$\frac{\partial}{\partial x_i} V(x) = \frac{\partial}{\partial x_i} J_*(\tilde{x})(1 - e^{-x_i/\gamma}), \qquad 1 \leq i \leq \ell,$$

and the right hand side is indeed zero when $x_i = 0$, proving (ii).

To see (iii) first consider a scaled state variable whose ith component is defined as

$$\tilde{x}_i^n = n^{-1}\big(nx_i + \gamma(e^{-nx_i/\gamma} - 1)\big)$$

We have,

$$V(nx) = J_*(n\tilde{x}^n) = n^2 J_*(\tilde{x}^n)$$

Obviously $\tilde{x}^n \to x$ as $n \to \infty$ for each $x \in \mathbb{R}_+^n$. Continuity of J_* completes the proof of (iii). ☐

FSAO can be established for either of the two myopic policies:

Theorem 6.3. *The following hold under either of the two feedback laws f_V or f_V^0 based on the function $V(x) = J_*(\tilde{x})$,*

(i) *There exists $K_{6.3} < \infty$ such that for each $x \in \mathbb{N}^\ell$,*

$$\mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x] \le -c(x) + K_{6.3} \log(1 + \|x\|).$$

(ii) *The policy is FSAO in the sense of (3).*

Proof. For simplicity we provide a proof only for f_V^0.

Writing $\widetilde{Q}(k) = \tilde{x}$ when $Q(k) = x$ we have by convexity of J_*,

$$\begin{aligned} J_*(\widetilde{Q}(k+1)) - J_*(\widetilde{Q}(k)) &\le \langle \nabla J_*(\widetilde{Q}(k+1)), \widetilde{Q}(k+1) - \widetilde{Q}(k) \rangle \\ &\le \langle \nabla J_*(\widetilde{Q}(k)), \widetilde{Q}(k+1) - \widetilde{Q}(k) \rangle \\ &\qquad + \|\nabla J_*(\widetilde{Q}(k+1)) - \nabla J_*(\widetilde{Q}(k))\| \|\widetilde{Q}(k+1) - \widetilde{Q}(k)\| \end{aligned}$$

Since J_* is assumed C^1 and has quadratic growth, it follows that ∇J is Lipschitz continuous. It then follows from the definition of V that for some constant $K_0 < \infty$, and any x, u,

$$\mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x, U(k) = u] \le \langle \nabla J_*(\tilde{x}), \widetilde{\Delta}(x,u) \rangle + K_0, \tag{66}$$

where $\widetilde{\Delta}(x,u) := \mathsf{E}[\widetilde{Q}(k+1) - \widetilde{Q}(k) \mid Q(k) = x, U(k) = u]$.

To bound the right hand side of (66) we compare with a new state variable that is more easily analyzed. Fix $\beta \ge 3\gamma$ and define for $x \in \mathbb{R}_+^\ell$,

$$x_i^0 = \begin{cases} x_i & \text{if } x_i \ge \beta \log(1 + \|x\|) \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{x}_i^0 = \begin{cases} \tilde{x}_i & \text{if } x_i \ge \beta \log(1 + \|x\|) \\ 0 & \text{otherwise,} \end{cases} \qquad i = 1, \dots, \ell.$$

In the analysis that follows we shall exploit the fact that $\tilde{x}^0 - x^0$ is very small: We have $|\tilde{x}_i^0 - x_i^0| \le \gamma e^{-x_i^0/\gamma}$ whenever the difference is non-zero, and this combined with the bound on β gives,

$$|\tilde{x}_i^0 - x_i^0| \le \gamma \exp\bigl(-(\beta/\gamma) \log(1 + \|x\|)\bigr) \le \gamma (1 + \|x\|)^{-3}, \qquad i = 1, \dots, \ell. \tag{67}$$

We now return to the inequality (66). To obtain an upper bound on the right hand side under the myopic policy f_V^0 we restrict to $u \in \mathsf{U}_0(x^0)$. Thus $u_i = 0$ whenever $x_i < \beta \log(1 + \|x\|)$. The following bounds are easily obtained under this assumption: we can find $K_1 < \infty$ such that with $x := Q(k)$ and $u \in \mathsf{U}_0(x^0)$,

$$\begin{aligned} \widetilde{Q}_i(k+1) - \widetilde{Q}_i(k) &\le Q_i(k+1) - Q_i(k), && x_i^0 = 0, \\ \bigl| \bigl(\widetilde{Q}_i(k+1) - \widetilde{Q}_i(k)\bigr) - \bigl(Q_i(k+1) - Q_i(k)\bigr) \bigr| &\le K_1 (1 + \|x\|)^{-3}, && x_i^0 = x_i, \end{aligned} \tag{68}$$

The first bound follows from the fact that $Q_i(k+1) \ge Q_i(k)$ a.s. if $U_i(k) = 0$, and the second bound is a consequence of (67).

Using (68) combined with monotonicity of J_*, and Lipschitz continuity of ∇J_*, we obtain for a possibly larger constant K_1,

$$\langle \nabla J_*(\widetilde{Q}(k)), \widetilde{Q}(k+1) - \widetilde{Q}(k) \rangle \le \langle \nabla J_*(\widetilde{Q}(k)), Q(k+1) - Q(k) \rangle + K_1 (1 + \|x\|)^{-2}.$$

This combined with the inequality (66) gives, for any $u \in \mathsf{U}_0(x^0)$,

$$\begin{aligned}\mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x, U(k) = u]\\ \le \langle \nabla J_*(\tilde{x}), Bu + \alpha \rangle + K_0 + K_1(1 + \|x\|)^{-2}\end{aligned}$$

Finally, applying once again Lipschitz continuity of the gradient ∇J_* gives for some $K_2 < \infty$,

$$\begin{aligned}\mathsf{E}[V(Q(k+1)) - V(Q(k)) \mid Q(k) = x, U(k) = u]\\ \le \langle \nabla J_*(x^0), Bu + \alpha \rangle + K_0 + K_1(1 + \|x\|)^{-2} + K_2\|\tilde{x} - x^0\|.\end{aligned}$$

The dynamic programming equation (8) for the fluid model gives $\min_{u \in \mathsf{U}_0(x^0)} \langle \nabla J_*(x^0), Bu + \alpha \rangle = -c(x^0)$, which implies that (i) holds for the policy f_V^0.

The proof of part (ii) is identical to the proof of Theorem 2.2 (i) using Lemma 6.2 (iii). □

7 Conclusions

This paper is the first to establish HTAO using a fixed policy when the cost is linear, and has provided an elementary approach to FSAO based on the fluid value function combined with a state transformation.

Many extensions are suggested by the results obtained here. It is likely that a completely general statement can be made for a wide class of network models, even when there are multiple bottlenecks, provided the effective cost of [29] is monotone. One possible approach to establish HTAO is to consider a multistep drift $P^T V - V$ for a fixed but large $T \ge 1$, rather than the one step drift condition (11). We conjecture that bounds analogous to (P4) may be obtained by combining some of the methods introduced here with the approach of [29, Section 4.2].

The story appears to be significantly more complex when the effective cost is not monotone since the optimal policy for the relaxation is not non-idling and hence not easily identified exactly [5]. In addition, it is shown in [5] that the 'gap' between the optimal policies for the fluid and stochastic models grows as $(1 - \rho_\bullet)^{-1}$, rather than the logarithm. However, given an optimal policy for a low-dimensional relaxation obtained numerically, it is possible that the methods of Section 6 can be used to obtain a HTAO solution for the ℓ-dimensional network.

The high sensitivity of cost with respect to safety-stock parameters shown in Figure 5 suggests that on-line learning approaches may be successfully implemented for performance improvement [17, 35]. The analysis of these algorithms will likely rely on parameterized Lyapunov functions of the form constructed in this paper.

Acknowledgment. This paper grew out of discussions with Professors Arie Hordijk, Peter Glynn, David McDonald and Ad Ridder at the Applied Probability Workshop held at the Mathematisches Forschungsinstitut Oberwolfach during the first week of December in 2003 (`http://www.mfo.de/`.) The paper was greatly improved following feedback from Profs. Jim Dai, Michael Veatch, and the anonymous reviewers. The author would like to express his sincere thanks to these colleagues for their comments and inspiration. He also expresses his thanks to the organizers of the Applied Probability Workshop and to MFO for creating this stimulating environment.

A Appendix: Bounds for the simple queue

In each result below we consider the CRW model (5) with $m^2 := \mathsf{E}[(A(1)-S(1))^2] < \infty$ and $\rho = \alpha/\mu < 1$. The policy is assumed non-idling: $U(k) = \mathbf{1}(Q(k) \ge 1)$. We begin with the

Proof of Proposition 1.1. We have from the definitions we have when $Q(0) = x$,

$$\begin{aligned} \mathsf{E}_x[Q(1)] &= x - \mu\mathbf{1}(x \neq 0) + \alpha \\ \mathsf{E}_x[Q(1)^2] &= x^2 - 2(\mu-\alpha)x + \mathbf{1}(x \neq 0)(m^2 - m_A^2) + m_A^2, \qquad x \in \mathbb{N}. \end{aligned} \tag{69}$$

The second identity implies a version of the drift inequality (V3) of [31]. It follows from the f-Norm Ergodic Theorem of [31] that a unique invariant probability measure exists with finite mean. A solution to Poisson's equation also exists which is bounded by a quadratic function of x (see [12, Theorem 2.3] and also [31, Chapter 17].)

The function h_* defined in (ii) is a quadratic function of the form $h_*(x) = ax^2 + bx$, with

$$a = \frac{1}{2}\frac{1}{\mu-\alpha}\Bigl(\frac{m^2 - m_A^2}{\mu}\Bigr), \quad b = \frac{1}{2}\frac{1}{\mu-\alpha}.$$

Writing $u = \mathbf{1}(x \neq 1)$ then gives,

$$\begin{aligned} \mathsf{E}_x[h_*(Q(1))]] &= h_*(x) + a[-\mu u + \alpha] + b[-2(\mu u - \alpha)x + u m^2 + (1-u)m_A^2] \\ &= h_*(x) - 2b(\mu u - \alpha)x + a[-\mu + \alpha] + bm^2, \end{aligned} \tag{70}$$

where we have used the fact that the precise values of a and b result in the cancellation,

$$a\mu + b[-m^2 + m_A^2] = 0.$$

We have $xu = x$ under the non-idling policy, so that the following identity is a consequence of (70):

$$\mathsf{E}_x[h_*(Q(1))] = h_*(x) - x + \tfrac{1}{2}\frac{\mu^{-1}(\alpha-\mu)(m^2 - m_A^2) + m^2}{\mu - \alpha}, \qquad x \in \mathbb{N}.$$

This is precisely Poisson's equation. It thus follows that h_* is the unique solution satisfying $h_*(0) = 0$, and η_* as defined in (i) is the steady state mean of $Q(k)$. □

Through similar arguments we obtain point-wise bounds on π.

Lemma A.1. *The unique invariant probability distribution π satisfies $\pi\{0\} = 1 - \rho$, and*

$$\pi\{n\} \le (\mu - \alpha)\frac{\pi\{[n+1,\infty)\}}{\mathsf{E}[(A(1) - S(1))_+]}, \qquad n \ge 1.$$

Proof. We have seen in Proposition 1.1 that π exists, with a finite, computable mean. Define for $x, n \in \mathbb{N}$,

$$u_n(x) = (x-n)_+, \quad g(n) = \mathsf{E}[(A(1) - S(1) - n)_+], \quad g_0(n) = \mathsf{E}[(A(1) - n)_+].$$

We have $\mathsf{E}[u_0(Q(k+1)) \mid Q(k) = x] = u_0(x) - (\mu - \alpha) + \mu\mathbf{1}_0(x)$. Stationarity of π then gives $\mu\pi\{0\} = (\mu - \alpha)$.

We have for each $x \in \mathbb{N}$, $k \geq 0$, $n \geq 1$,

$$\mathsf{E}[u_n(Q(k+1)) \mid Q(k) = x] = u_n(x) + g(n-x)\mathbf{1}_{[1,n]}(x) + g_0(n)\mathbf{1}_0(x) - (\mu - \alpha)\mathbf{1}_{[n+1,\infty]}(x),$$

which together with stationarity of π gives the identitiy,

$$\pi(0)g_0(n) + \sum_{k=1}^{n} \pi(k)g(n-k) = (\mu - \alpha) \sum_{k=n+1}^{\infty} \pi(k).$$

This implies the desired bound since the left-hand side is no less than $\pi(n)g(0)$ for each $n \geq 1$. □

Lemma A.2. *The first bound below holds when* $\boldsymbol{Q}$ *is stationary, with* $Q(0) \sim \pi$*. The second holds for any initial condition* $x \in \mathbb{N}$*, and any stopping time* τ*:*

$$\begin{aligned} \mathsf{E}_\pi\bigl[\bigl(1+Q(0)\bigr)^{-2}\bigr] &\leq \frac{2(\mu-\alpha)}{\mathsf{E}[(A(1)-S(1))_+]}; \\ \mathsf{E}_x\Bigl[\sum_{k=0}^{\tau-1}\bigl(1+Q(k)\bigr)^{-2}\Bigr] &\leq \mathsf{E}_x[\tau] + (\mu-\alpha)^{-1}\mathsf{E}_x\Bigl[Q(\tau) + \bigl(1+Q(\tau)\bigr)^{-1} - [x + (1+x)^{-1}]\Bigr]. \end{aligned}$$

Proof. Lemma A.1 provides the first bound:

$$\mathsf{E}_\pi\bigl[\bigl(1+Q(0)\bigr)^{-2}\bigr] \leq (\mu-\alpha)\frac{1}{\mathsf{E}[(A(1)-S(1))_+]}\sum_{k=0}^{\infty}(1+k)^{-2}$$

To see the second bound we use convexity of the function $(1+x)^{-1}$, $x \geq 0$, which provides the following lower bound,

$$\begin{aligned} &\bigl(1+Q(k+1)\bigr)^{-1} \\ &\quad\geq \bigl(1+Q(k)\bigr)^{-1} - \bigl(1+Q(k)\bigr)^{-2}\bigl(Q(k+1)-Q(k)\bigr) \\ &\quad= \bigl(1+Q(k)\bigr)^{-1} - \bigl(1+Q(k)\bigr)^{-2}\bigl(A(k+1)-S(k+1)\bigr) - \mathbf{1}(Q(k)=0)S(k+1), \end{aligned}$$

where we have used the recursive formula,

$$Q(k+1) = Q(k) + \bigl(A(k+1) - S(k+1)\bigr) + \mathbf{1}(Q(k)=0)S(k+1), \qquad k \geq 0.$$

This identity can be used once more to eliminate the indicator in the previous bound, giving

$$\Bigl[\bigl(1+Q(k+1)\bigr)^{-1}+Q(k+1)\Bigr] \geq \Bigl[\bigl(1+Q(k)\bigr)^{-1}+Q(k)\Bigr] + \bigl(1-\bigl(1+Q(k)\bigr)^{-2}\bigr)\bigl(A(k+1)-S(k+1)\bigr).$$

Summing over k from 0 to $\tau - 1$ and taking expectations then gives,

$$\mathsf{E}_x\Bigl[Q(\tau) + \bigl(1+Q(\tau)\bigr)^{-1} - [x+(1+x)^{-1}]\Bigr] \geq (\mu-\alpha)\Bigl(\mathsf{E}_x\Bigl[\sum_{k=0}^{\tau-1}\bigl(1+Q(k)\bigr)^{-2}\Bigr] - \mathsf{E}_x[\tau]\Bigr).$$

□

Lemma A.3. *The following bounds hold for any non-zero initial condition $x \in \mathbb{N}$, with τ_0 equal to the first hitting time to the origin:*

$$\mathsf{E}_x\big[\tau_0\big] = \frac{x}{\mu - \alpha}, \qquad \mathsf{Var}_x\big[\tau_0\big] \le m^2 \frac{x}{(\mu-\alpha)^3},$$

where $m^2 = \mathsf{E}[(S(1) - A(1))^2]$.

Proof. Consider the two test functions,

$$V_1(x) = \frac{x}{\mu-\alpha}, \quad V_2(x) = \tfrac{1}{2}\frac{x^2}{\mu-\alpha}, \qquad x \in \mathbb{N},$$

that satisfy the identities,

$$PV_1\,(x) = V_1(x) - 1, \quad PV_2\,(x) = V_2(x) - x + \tfrac{1}{2}\frac{m^2}{\mu-\alpha}, \qquad x \ge 1.$$

The first identity implies that $M_1(k) := k \wedge \tau_0 + V_1(Q(k \wedge \tau_0))$, $k \ge 0$, is a martingale. The second identity combined with the Comparison Theorem of [31] gives,

$$\mathsf{E}_x\Big[\sum_{k=0}^{\tau_0-1} Q(k)\Big] \le V_2(x) + \tfrac{1}{2}\frac{m^2}{\mu-\alpha}\mathsf{E}_x\big[\tau_0\big], \qquad x \ge 1. \tag{71}$$

This bound implies that $\boldsymbol{M}_1$ is uniformly integrable, and this implies the formula for the mean hitting time to the origin.

To see the variance bound we again apply (71). The sum on the left hand side of (71) has the following representation:

$$\begin{aligned}
\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\tau_0-1} Q(k)\Big] &= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\tau_0-1} V_1(Q(k))\Big] \\
&= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\tau_0-1} \mathsf{E}_{Q(k)}[\tau_0]\Big] \\
&= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\infty} \mathsf{E}_{Q(k)}[\tau_0]\mathbf{1}(k < \tau_0)\Big] \\
&= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\infty} (\tau_0 - k)\mathbf{1}(k < \tau_0)\Big],
\end{aligned}$$

where the last identity follows from the the smoothing property of the conditional expectation

$$\mathsf{E}_{Q(k)}[\tau_0] = \mathsf{E}[\tau_0 - k \mid Q(0), \dots, Q(k)], \quad k \ge 0 \text{ on } \{\tau_0 > k\}.$$

The right hand side may be further transformed as follows,

$$\begin{aligned}
\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\tau_0-1} Q(k)\Big] &= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\infty} (\tau_0 - k)\mathbf{1}(k < \tau_0)\Big] \\
&= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=0}^{\tau_0-1} (\tau_0 - k)\Big] \\
&= (\mu-\alpha)\mathsf{E}_x\Big[\textstyle\sum_{k=1}^{\tau_0} k\Big] \\
&= \tfrac{1}{2}(\mu-\alpha)\mathsf{E}_x\big[\tau_0^2 + \tau_0\big].
\end{aligned}$$

Combining this with (71) then gives

$$\mathsf{E}_x\big[\tau_0^2\big] \le \frac{2}{\mu-\alpha}\big(V_2(x) + \tfrac{1}{2}\frac{m^2}{\mu-\alpha}\mathsf{E}_x\big[\tau_0\big]\big),$$

and substituting the expressions for $V_2(x)$ and $\mathsf{E}_x\big[\tau_0\big]$ then gives,

$$\mathsf{E}_x\big[\tau_0^2\big] \le \Big(\frac{x}{\mu-\alpha}\Big)^2 + m^2\frac{x}{(\mu-\alpha)^3}.$$

We finally arrive at the desired bound from the identity $\mathsf{Var}_x\big[\tau_0\big] = \mathsf{E}_x\big[\tau_0^2\big] - \mathsf{E}_x\big[\tau_0\big]^2$. □

B Appendix: Bounds for a parameterized family

Here we provide a proof of Theorem 2.2.

It is assumed throughout this section that the family of network models described in Section 2 satisfies (P1)–(P4). In the results below we suppress dependency of $\boldsymbol{W} = \boldsymbol{W}^\vartheta$ and $\boldsymbol{Q} = \boldsymbol{Q}^\vartheta$ on the parameter ϑ whenever possible to simplify notation.

For $n \ge 1$, $x \in \mathbb{R}_+^\ell$, recall from (2) the definition of $J_n(x,T)$. A lower bound is easily obtained (see [27, 26] for related results).

Lemma B.1. *The lower bound holds,*

$$\liminf_{n\to\infty} J_n(x,T) \ge J_*(x), \qquad T \ge T_\circ(x)\,,$$

where $T_\circ(x)$ denotes the draining time for the optimal fluid model with initial condition $x \in \mathbb{R}_+^\ell$.

Proof. The piece-wise linear function of time $\{\mathsf{E}[q_n(t;x)] : t \ge 0\}$ is a feasible trajectory for the fluid model, with initial condition $y^n = n^{-1}[nx]$. Consequently, since the cost is linear,

$$\mathsf{E}\Big[\int_0^T c(q_n(t;x))\,dt\Big] \ge \inf \int_0^T c(q(t;y^n))\,dt,$$

where the infimum is over all feasible fluid trajectories. The right hand side is equal to $J_*(y^n)$ for $T \ge T_\circ(y^n)$. Continuity of J_* completes the proof. □

Under (P1)–(P4) we have a version of the drift inequality (V3) of [31]:

$$PV\,(x) \le V(x) - \tfrac{1}{2}c(x) + \widehat{\eta}_* + K_1\big(\mathbf{1}_S(x) + (1-\vartheta)^{-1}(1+w_j(x))^{-2}\big), \qquad x \in \mathbb{N}^\ell, \qquad (72)$$

where $w_j(x) = \langle \xi^j, x\rangle$, K_1 is a fixed constant, and $S \subset \mathbb{N}^\ell$ is a fixed finite set, both independent of ϑ. We obtain the following crude upper bounds based on this drift inequality:

Lemma B.2. *The following bounds hold for some fixed $K_{\rm B.2} < \infty$, and all $x \in \mathbb{N}^\ell$, $T > 0$, $\vartheta \in [0,1)$:*

$$\begin{aligned} J_n(x,T) &\le K_{\rm B.2}\,(1-\vartheta)^{-1}\Big(\|x\|^2 + n^{-1}T\Big); \\ \eta := \mathsf{E}_\pi[c(Q(k))] &\le K_{\rm B.2}\,(1-\vartheta)^{-1}. \end{aligned}$$

Proof. By the Comparison Theorem of [31] and the drift inequality (72) it follows that for any stopping time τ and any initial condition,

$$\tfrac{1}{2}\mathsf{E}_x\Big[\sum_{k=0}^{\tau-1} c((Q(k)))\Big] \le V(x) + \mathsf{E}_x\Big[\sum_{k=0}^{\tau-1}\big(\widehat{\eta}_* + K_1(1+(1-\vartheta)^{-1}\big)\Big].$$

On setting $\tau = nT$ for a given $T > 0$, $n \geq 1$, we obtain for all $x \in \mathbb{R}^\ell$,

$$\tfrac{1}{2} J_n(x,T) \leq n^{-2}\Big(V([nx]) + nT\big(\widehat{\eta}_* + K_1(1+(1-\vartheta)^{-1})\big)\Big),$$

and the first bound follows from the bound on V assumed in (P4).

The second bound follows on setting $\tau = N$ for $N \geq 1$, giving

$$\frac{1}{N}\tfrac{1}{2}\mathsf{E}_x\Big[\sum_{k=0}^{N-1} c((Q(k)))\Big] \leq \frac{1}{N}V(x) + \widehat{\eta}_* + K_1\big(1+(1-\vartheta)^{-1}\big).$$

Letting $N \to \infty$ we obtain the bound $\eta \leq 2\big(\widehat{\eta}_* + K_1(1+(1-\vartheta)^{-1})\big)$. □

A sharp bound on the average cost follows from these relatively crude bounds:

Proof of Theorem 2.2 (ii). The formula (27) follows directly from Proposition 1.1 and the identity $\lambda_j/\mu_j = \vartheta$ for $\widehat{\boldsymbol{W}}^*_j$.

Recall that $\{W_j(k), \widehat{W}^*_j(k) : k \geq 0\}$ are defined on the same probability space, and have a jointly stationary version. We consider this stationary process in the following.

We now apply the bound (29) obtained using the Comparison Theorem. This combined with (P4) gives,

$$\eta \leq \widehat{\eta}_* + K_{\rm P4}\,\mathsf{E}_\pi\Big[\log(1+c(Q(k))) + (1-\vartheta)^{-1}\big(1+W_j(k)\big)^{-2}\Big]. \tag{73}$$

Jensen's inequality combined with Lemma B.2 provides a bound on the first term within the expectation,

$$\mathsf{E}_\pi[\log(1+c(Q(k)))] \leq \log(1+\mathsf{E}_\pi[c(Q(k))]) \leq \log(1+K_{\rm B.2}\,(1-\vartheta)^{-1}).$$

Minimality of the relaxation implies the following bound on the second term:

$$\mathsf{E}_\pi\big[\big(1+W_j(k)\big)^{-2}\big] \leq \mathsf{E}_\pi\big[\big(1+\widehat{W}^*_j(k)\big)^{-2}\big].$$

Consequently, applying Lemma A.2 (i) to $\widehat{\boldsymbol{W}}^*_j$, it is apparent that (73) gives the desired bound on η provided the denominator $\mathsf{E}[(L_j(1) - S_j(1))_+]$ appearing in an application of the lemma is bounded away from zero for ϑ in a neighborhood of $\vartheta = 1$.

Note that $\mathsf{E}[(L_j(1) - S_j(1))_+] \neq 0$ when $\vartheta = 1$ by the assumption that $\mathsf{E}[L_j(1) - S_j(1)] = 0$ combined with the irreducibility assumption (P3). The monotonicity assumption (P2) implies that $\mathsf{E}[(L_j(1) - S_j(1))_+] \neq 0$ for $\vartheta \sim 1$. □

To refine the upper bound on J_n we require the bounds obtained in the following two lemmas:

Lemma B.3. *There exists $K_{\rm B.3} < \infty$ such that for each non-zero $x \in \mathbb{N}^\ell$ and $\vartheta_0 \leq \vartheta < 1$,*

$$\mathsf{E}_x\Big[\sum_{k=0}^{\widehat{T}_*(x)-1}(1+W_j(k))^{-2}\Big] \leq \mathsf{E}_x\Big[\sum_{k=0}^{\widehat{T}_*(x)-1}(1+\widehat{W}^*_j(k))^{-2}\Big] \leq K_{\rm B.3}\,\frac{\sqrt{\|x\|}}{(1-\vartheta)^{5/2}},$$

where $\widehat{T}_*(x) := (\mu_j - \lambda_j)^{-1}\langle \xi^j, x\rangle$.

Proof. The first inequality follows from minimality of $\widehat{\boldsymbol{W}}_j^*$.

To see the second bound we apply Lemma A.2 to $\widehat{\boldsymbol{W}}_j^*$ to obtain, for any $T \ge 0$,

$$\mathsf{E}_x\Big[\sum_{k=0}^{T-1}(1+\widehat{W}_j^*(k))^{-2}\Big] \le T + (\mu_j - \lambda_j)^{-1}\big(\mathsf{E}_x[\widehat{W}_j^*(T)] + 1 - w\big), \tag{74}$$

where $w = \widehat{W}_j^*(0) = \langle \xi^j, x\rangle$. For $T = \widehat{T}_*(x)$, the right hand side becomes,

$$(\mu_j - \lambda_j)^{-1}(\mathsf{E}_x[\widehat{W}_j^*(T)] + 1).$$

This is bounded as follows: For any $T \ge 0$,

$$\mathsf{E}_x[\widehat{W}_j^*(T)] \le w - (\mu_j - \lambda_j)T + \mathsf{E}_x[(T - \hat{\tau}_0)_+]$$

where $\hat{\tau}_0$ denotes the first hitting time to the origin for $\widehat{\boldsymbol{W}}_j^*$. For $T = \widehat{T}_*(x)$ we apply Jensen's inequality and Lemma A.3 to obtain,

$$\mathsf{E}_x[\widehat{W}_j^*(T)] \le \mathsf{E}_x[(T-\hat{\tau}_0)_+] \le \sqrt{\mathsf{E}_x[(T-\hat{\tau}_0)^2]} \le \sqrt{\hat{m}^2 \frac{w}{(\mu_j - \lambda_j)^3}}, \tag{75}$$

where $\hat{m}^2$ is the variability parameter for $\widehat{\boldsymbol{W}}_j^*$ defined in Lemma A.3.

Combining (74) and (75) with $T = \widehat{T}_*(x)$ gives the desired bound. □

Lemma B.4. *For any $T > 0$, $n \ge 1$, $x \in \mathbb{R}_+^\ell$,*

$$\mathsf{E}_{[nx]}\Big[\sum_{k=0}^{nT-1}\log(1 + c(Q(k))\Big] \le nT\log\Big(1 + T^{-1}nJ_n(x,T)\Big).$$

Proof. From Jensen's inequality we have,

$$\frac{1}{nT}\mathsf{E}_{[nx]}\Big[\sum_{k=0}^{nT-1}\log(1+c(Q(k))\Big] \le \log\Big(1 + \frac{1}{nT}\mathsf{E}_{[nx]}\Big[\sum_{k=0}^{nT-1} c(Q(k))\Big]\Big).$$

The result then follows from the definition of $J_n(x,T)$. □

Proof of Theorem 2.2 (i). The lower bound is given in Lemma B.1.

For the upper bound we again apply the Comparison Theorem to (11), which gives the following bound with $T = T_\circ = O(\|x\|(1-\vartheta)^{-1})$, $x \in \mathbb{R}_+^\ell$, and $n \ge 1$:

$$\begin{aligned} J_n(x,T) \le n^{-2}V([nx]) + n^{-1}T\widehat{\eta}_* + n^{-2}K_{\text{P4}}\frac{1}{1-\vartheta}\mathsf{E}_{[nx]}\Big[\sum_{k=0}^{nT-1}(1+W_j(k))^{-2}\Big] \\ + n^{-2}K_{\text{P4}}\mathsf{E}_{[nx]}\Big[\sum_{k=0}^{nT-1}\log(1+c(Q(k))\|\Big]. \end{aligned}$$

The two expectations on the right hand side may be bounded using Lemma B.3 and Lemma B.4, respectively, yielding,

$$\limsup_{n\to\infty}\frac{n}{\log(n)}\Big(J_n(x,T) - n^{-2}V([nx])\Big) \le \limsup_{n\to\infty}\frac{1}{\log(n)}T\log\Big(1+T^{-1}nJ_n(x,T)\Big) \le T, \tag{76}$$

where the final inequality follows from Lemma B.2. Note that one can establish the bound $T_\circ(x) - \widehat{T}_*(x) \le K\|x\|$, $x \in \mathbb{N}^\ell$, where the constant K is independent of load, which justifies the application of Lemma B.3 in the bound above.

Moreover, it follows from (P4) and the fact that J_* is piecewise-quadratic that

$$0 \le \limsup_{n\to\infty} \frac{n}{\log(n)} \Big(n^{-2} V([nx]) - J_*(x) \Big) \le K_{\rm P4} \frac{\|x\|}{1-\vartheta}, \qquad \vartheta \in [\vartheta_0, 1). \tag{77}$$

In particular, we see that $\lim_{n\to\infty} n^{-2}V([nx]) = J_*(x)$.

Finally, we combine (76) and (77) to obtain the desired bound:

$$\begin{aligned} \limsup_{n\to\infty} \frac{n}{\log(n)} \Big(J_n(x,T) - J_*(x) \Big) \;\le\; & \limsup_{n\to\infty} \frac{n}{\log(n)} \Big(n^{-2} V([nx]) - J_*(x) \Big) \\ & + \limsup_{n\to\infty} \frac{n}{\log(n)} \Big(J_n(x,T) - n^{-2} V([nx]) \Big) \\ \le\; & K_{\rm P4} \frac{\|x\|}{1-\vartheta} + T. \end{aligned}$$

This completes the proof since we are taking $T = T_\circ = O(\|x\|(1-\vartheta)^{-1})$. □

References

[1] Nicole Bäuerle. Optimal control of queueing networks: an approach via fluid models. *Adv. in Appl. Probab.*, 34(2):313–328, 2002.

[2] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with complete resource pooling: Asymptotic optimality of a continuous review threshold policy. *Ann. Appl. Probab.*, 11:608–649, 2001.

[3] M. Bramson. State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems: Theory and Applications*, 30:89–148, 1998.

[4] H. Chen and D. D. Yao. *Fundamentals of queueing networks: Performance, asymptotics, and optimization.* Springer-Verlag, New York, 2001. Stochastic Modelling and Applied Probability.

[5] M. Chen, C. Pandit, and S. P. Meyn. In search of sensitivity in network optimization. *Queueing Systems: Theory and Applications*, 44(4):313–363, 2003.

[6] 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.

[7] J. G. Dai and W. Lin. Maximum pressure policies in stochastic processing networks. Operations Research Volume 53, Number 2, to appear.

[8] J. G. Dai and Y. Wang. Nonexistence of Brownian models of certain multiclass queueing networks. *Queueing Systems: Theory and Applications*, 13:41–46, May 1993.

[9] A. Ephremides and L. Tassiulas. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automat. Control*, 37:1936–1949, 1992.

[10] A. Gajrat, A. Hordijk, and A. Ridder. Large deviations analysis of the fluid approximation for a controllable tandem queue. *Ann. Appl. Probab.*, 13:14231448, 2003.

[11] L. Georgiadis, W. Szpankowski, and L. Tassiulas. A scheduling policy with maximal stability region for ring networks with spatial reuse. *Queueing Systems: Theory and Applications*, 19(1-2):131–148, 1995.

[12] P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.

[13] L. Green. A queueing system with general-use and limited-use servers. *Oper. Res.*, 33(1):168–182, 1985.

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

[15] J. M. Harrison. *Brownian motion and stochastic flow systems*. Robert E. Krieger Publishing Co. Inc., Malabar, FL, 1990. Reprint of the 1985 original.

[16] J. M. Harrison and M.J. López. Heavy traffic resource pooling in parallel-server systems. *Queueing Systems: Theory and Applications*, 33:339–368, 1999.

[17] S. G. Henderson, S. P. Meyn, and V. B. Tadić. Performance evaluation and policy selection in multiclass networks. *Discrete Event Dynamic Systems: Theory and Applications*, 13(1-2):149–189, 2003. Special issue on learning, optimization and decision making (invited).

[18] 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.

[19] L. Kleinrock. *Queueing Systems Volume 1: Theory*. Wiley, 1975.

[20] G. Koole. On the pathwise optimal Bernoulli routing policy for homogeneous parallel servers. *Math. Operations Res.*, 21:469–476, 1996.

[21] V. G. Kulkarni. *Modeling and analysis of stochastic systems*. Texts in Statistical Science Series. Chapman and Hall Ltd., London, 1995.

[22] H. J. Kushner. *Heavy traffic analysis of controlled queueing and communication networks*. Springer-Verlag, New York, 2001. Stochastic Modelling and Applied Probability.

[23] N. Laws. *Dynamic routing in queueing networks*. PhD thesis, Cambridge University, Cambridge, UK, 1990.

[24] C. Maglaras. Dynamic scheduling in multiclass queueing networks: Stability under discrete-review policies. *Queueing Systems: Theory and Applications*, 31:171–206, 1999.

[25] A. Mandelbaum and A. L. Stolyar. Scheduling flexible servers with convex delay costs: heavy-traffic optimality of the generalized $c\mu$-rule. *Operations Res.*, 52(6):836–855, 2004.

[26] S. P. Meyn. The policy iteration algorithm for average reward Markov decision processes with general state space. *IEEE Trans. Automat. Control*, 42(12):1663–1680, 1997.

[27] S. P. Meyn. Stability and optimization of queueing networks and their fluid models. In *Mathematics of stochastic manufacturing systems (Williamsburg, VA, 1996)*, pages 175–199. Amer. Math. Soc., Providence, RI, 1997.

[28] S. P. Meyn. Sequencing and routing in multiclass queueing networks. Part I: Feedback regulation. *SIAM J. Control Optim.*, 40(3):741–776, 2001.

[29] S. P. Meyn. Sequencing and routing in multiclass queueing networks. Part II: Workload relaxations. *SIAM J. Control Optim.*, 42(1):178–217, 2003.

[30] S. P. Meyn. Value functions, optimization, and performance evaluation in stochastic network models. Submitted for publication, 2003.

[31] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993. online: `http://black.csl.uiuc.edu/~meyn/pages/book.html`.

[32] T. Osogami, A. Harchol-Balter, M. Scheller-Wolf, and L. Zhang. Exploring threshold-base policies for load sharing. In *Proceedings of the Forty-second Annual Allerton Conference on Communication, Control, and Computing*, 2004.

[33] S. Stidham, Jr. and R. Weber. A survey of Markov decision models for control of networks of queues. *Queueing Systems: Theory and Applications*, 13(1-3):291–314, 1993.

[34] A. L. Stolyar. Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. *Adv. Appl. Probab.*, 14(1):1–53, 2004.

[35] V. Tadic and S.P. Meyn. Asymptotic properties of two time-scale stochastic approximation algorithms with constant step sizes. In *American Control Conference*, 2003.

[36] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automat. Control*, 37(12):1936–1948, 1992.

[37] Y-C Teh and A. R. Ward. Critical thresholds for dynamic routing in queueing networks. *Queueing Systems: Theory and Applications*, 42(3):297–316, 2002.

[38] M. H. Veatch. Fluid analysis of arrival routing. *IEEE Trans. Automat. Control*, AC-46:1254–1257, 2001.

[39] G. Weiss. Optimal draining of a fluid re-entrant line. In Frank Kelly and Ruth Williams, editors, *Volume 71 of IMA volumes in Mathematics and its Applications*, pages 91–103, New York, 1995. Springer-Verlag.

[40] R. J. Williams. Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems: Theory and Applications*, 30(1-2):27–88, 1998.