Continuous-review tracking policies for dynamic control of stochastic networks

Constantinos Maglaras [a] *

[a] *Columbia Business School,*
Uris Hall 409, 3022 Broadway,
New York, NY 10027-6902
E-mail: c.maglaras@columbia.edu

Received January 2002; revised July 2002

This paper is concerned with dynamic control of stochastic processing networks. Specifically, it follows the so called "heavy traffic approach," where a Brownian approximating model is formulated, an associated Brownian optimal control problem is solved, the solution of which is then used to define an implementable policy for the original system. A major challenge is the step of policy translation from the Brownian to the discrete network. This paper addresses this problem by defining a general and easily implementable family of *continuous-review tracking* policies. Each such policy has the following structure: at each point in time t, the controller observes the current vector of queue lengths q and chooses (i) a target position $z(q)$ of where the system should be at some point in the near future, say at time $t+l$, and (ii) an allocation vector $v(q)$ that describes how to split the server's processing capacity amongst job classes in order to steer the state from q to $z(q)$. Implementation of such policies involves the enforcement of small safety stocks. In the context of the "heavy traffic" approach, the solution of the approximating Brownian control problem is used in selecting the target state $z(q)$. The proposed tracking policy is shown to be asymptotically optimal in the heavy traffic limiting regime, where the Brownian model approximation becomes valid, for multiclass queueing networks that admit *orthant* Brownian optimal controls; this is a form of pathwise, or greedy, optimality. Several extensions are discussed.

Keywords: Continuous-review, tracking policies, multiclass networks, fluid models, Brownian models, heavy-traffic, asymptotic optimality.

AMS Subject classification: Primary: 90B35, 60K25; Secondary: 93E15, 93D30.

1. Introduction

The manufacturing, service operations and communications industries offer many examples of technological systems in which *units of work* flow through a system of *processing resources* and have a sequence of *activities* performed on them (or by them), and in which the workflow is subject to stochastic variability. These three basic elements characterize a *stochastic processing network*. As an example,

* A preliminary version of this paper has been published in the Proc. of the 37th Allerton Conf. on Communication, Control and Computing, Univ. Illinois, Urbana-Champaign, 1999.

consider a semiconductor manufacturer that produces Application-Specific-Integration-Circuits (ASIC) on a contract basis. It is capable of producing many different types of products (limited only by its production technology), and each product requires several hundreds of processing steps performed at dozens of different machines.

Dynamic flow management capability in such systems takes the form of admission decisions at the edges of the network, routing of work through the network, and sequencing of activities at each resource. This paper is about optimal sequencing in such systems with respect to an expected weighted inventory cost. Multiclass queueing networks are natural mathematical models for such systems. While they provide a detailed description of system dynamics, with the exception of restricted examples, they are not amenable to exact analysis and optimization (see [39]). As a result, such problems are typically addressed using heuristics that are validated via simulation.

In contrast, this paper articulates a general family of *continuous-review, tracking* policies that are derived through a systematic approach, they are easy to describe, implement and analyze, and achieve near-optimal performance. Specifically, the proposed have a *receding -or rolling- horizon* structure that is outlined below. At any point in time t, a tracking policy reviews the state of the system, which will be denoted by q (mnemonic for the vector of queue lengths), and computes:

(i) a *target* state, $z(q)$, of where the system should be at some point in the near future at time $t + l$, where l is an appropriately selected moderate planning horizon; and

(ii) a nominal control, $v(q)$, that specifies how to split server effort amongst job classes (i.e., the different *activities*) in order to achieve the target $z(q)$ in l time units. [1]

The basic structure of these policies is borrowed from the area of model predictive control, where it has been applied very successfully in the control of nonlinear dynamical systems; see [12]. Continuous-review policies have also been used in the areas of forecasting or production, planning and control and in inventory management; see [41] and [2], respectively. This structure is general enough to describe almost any sequencing policy of interest. Instead of trying to find the optimal such rule that is analytically and numerically intractable, our focus will be on demonstrating how it can be used as a translation mechanism for policies derived via an analysis of a hierarchy of approximate models, and in particular fluid and diffusion approximations.

Fluid and diffusion (or Brownian) models provide tractable "relaxations" for the original systems under investigation, and offer a powerful framework for analysis and synthesis. The main focus of this paper will be on the so called "heavy traffic" approach that deals with stochastic networks where at least some of their resources are operating under nominal loads that are close to their respective capacities. For such systems, *one first approximates (or replaces) the stochastic network by its Brownian analog, then solves an associated Brownian optimal control problem (BCP), and finally, defines an*

[1] For example, if a server is responsible for activities 1 and 2 and $v(q) = [.7,\ .3]$, then the server should split its capacity and allocate 70% into processing activity 1 and the remaining 30% to activity 2. The control $v(q)$ stays constant between state transitions, where q, and potentially $v(q)$, changes.

implementable rule for the stochastic network by translating the associated Brownian optimal control policy. This heavy traffic approach was proposed by Harrison [19], and has been successfully applied in many examples [25,26,32,35]. A major obstacle in this approach is that the Brownian control policy does not have a direct interpretation in the original stochastic network. Proposed translations in the examples that appear in the papers mentioned above have exploited specific problem structure and are hard to generalize. Moreover, the heavy traffic asymptotic analysis under the implemented policies had proved to be very complex even in simple examples [32,35]. A summary of these issues dating in 1996 is given in [44].

Recent work has generalized the network models for which this approach is applicable (see [22,23, 37]). A general translation mechanism based on discrete-review policies has been advanced by Harrison [20]. Such policies review system status in discrete time intervals and make scheduling decisions over each of these intervals using the solution of the Brownian control problem. This structure was adopted by the author in proposing a family of trajectory tracking policies [33,34] based on fluid model analysis. The policies proposed in [34] are shown to be asymptotically optimal under fluid scaling –a form of transient optimality. A related result was obtained in [3]. The last few papers are related to a strand of work that uses fluid as opposed to Brownian system model for network analysis [14,9,6] and policy synthesis [10,1,36,33,34,37,38]. Recent work by Meyn [37,38] and Chen *et.al.* [11] explores connections between fluid and Brownian models in designing good policies for the underlying network models, and studies their asymptotic performance. Finally, in terms of asymptotic analysis, Bramson [7] and Williams [46,45] have advanced a powerful technique that reduces the diffusion scale analysis into a verification of fluid scale properties; see also Bramson and Dai [8] for a discussion and extension of this framework, and Kumar [29] for an application in proving the heavy traffic asymptotic optimality of a static priority rule proposed in [26] for two station closed networks.

The main focus of this paper is to describe a general translation mechanism in the form of continuous-review policies and provide a rigorous asymptotic analysis of their performance for a class of network control problems that admit a simple form of Brownian optimal control policy. The main contributions are the following.

1. In terms of the solution methodology, we show that the continuous-review structure sketched above, and described in detail in §4, provides a general translation mechanism from the optimal solution of the approximating Brownian model to an implementable policy for the stochastic network of original interest. Broadly speaking, the Brownian solution is used to define the target $z(q)$ and fluid reasoning is used to compute the appropriate control (sequencing decisions) $v(q)$ to achieve this target.

2. In terms of theoretical analysis, we establish that as the system approaches the heavy traffic regime, where the Brownian model approximation becomes valid, the "translated" continuous-review policy achieves asymptotically optimal performance. This result is proved for the class of problems that admit "orthant" Brownian optimal controls (defined in §3); this roughly corresponds to problems for which the greedy policy that myopically drains cost out of the system is optimal in the associated BCP formulation. This restriction is not necessary for the translation step, but is needed to push through the asymptotic analysis. Theorem 1 of this paper is one of the first such asymptotic optimality results that is not example-specific, but addresses a general class of problems.

3. Both the structure of the policy and its analysis yield some interesting insights and consequences. (i) We deduce that the continuous-review structure described here can also be used to translate policies extracted via a fluid model analysis in a way that achieves asymptotically optimal transient response. (ii) In addition, for the class of problems that admit "orthant" Brownian optimal controls, the continuous-review implementation of the associated optimal fluid control policy is shown to be asymptotically optimal in the heavy traffic regime. This implies that fluid and Brownian analysis are equivalent for this class of problems, where the "greedy" policy is optimal in the respective limiting control problems.

Points 2 and 3 are related to results that have been obtained by Meyn [38] in parallel work to ours. Specifically, restricting attention to problems that admit pathwise optimal controls, Meyn showed [38, Thm. 4.3-4.5] that a family of appropriately defined discrete-review policies is asymptotically optimal in the heavy traffic limit. His result seems to hold for a more general class of underlying network models. However, in both cases and while restricting attention to problems that admit pathwise Brownian optimal controls, the essential issues that need to be addressed in the derivation of the corresponding results are still the translation mechanism and the asymptotic analysis. In terms of these two issues, the respective results are different. Ours uses continuous-review policies analyzed using the framework developed by Bramson [7] and Williams [46,45], while [38] employs discrete-review implementations and a different style of analysis.

Teh [43] and Bell and Williams [4] have also proposed continuous-review -or more accurately threshold- policies based on Brownian model analysis. Teh constructed dynamic priority rules by shrinking the planning horizon of a discrete-review policy [20,34] down to zero; this method is applicable to multiclass networks with pathwise Brownian optimal controls, and their asymptotic analysis is example-specific. Bell and Williams [4] proposed a threshold policy for multiclass systems with complete resource pooling; the latter implies a one-dimensional associated Brownian model. Both [43,4] have a different flavor from our work that instead aims to give a generic definition of a continuous-review policy that is more widely applicable. From a practical viewpoint, continuous-review policies may be appealing in applications where server splitting is allowed and where it is more natural to think of the control in terms of fractional capacity allocations to different job classes. Also, these policies connect to other existing literature in stochastic modelling where the same form of a control appears [41,2,12], and as it has been numerically observed in [21] they are expected to lead to performance improvements over the existing discrete-review implementations when implemented in the original stochastic networks. While these performance gains are asymptotically negligible if one uses the correct form of a continuous- or discrete-review policy, they may still be of interest to the practitioner. Finally, the structure of continuous-review policies may be applicable as a control mechanism in complex, stochastic systems that arise in other areas of operations management and service operations, as in the optimization of production-inventory systems and in revenue management. In common to the problems studied here, detailed modelling in these areas often leads to intractable formulations that researchers have addressed using approximate analysis. The tracking structure, the choice of planning horizon l, and their asymptotic analysis and performance guarantees provide useful insights that can be ported in these application domains.

The remainder of the paper is structured as follows. §2 describes the network model, its fluid and Brownian analogues. §3 describes the class of "orthant" Brownian controls. §4 defines a tracking policy based on the solution of the associated Brownian control problem, and §5 states our main result regarding the heavy-traffic asymptotic optimality of this policy. §6 gives a discussion of this result and §7 extends the translation mechanism to problems whose associated Brownian models do not admit "orthant" Brownian optimal controls.

2. Network models

This section describes the detailed queueing network model and the control problem of original interest, its fluid and Brownian approximations, and the associated Brownian control problem.

2.1. Open multiclass queueing networks

Consider a queueing network of single server stations indexed by $i \in \mathcal{I} = \{1, \ldots, S\}$. The network is populated by job classes indexed by $k \in \mathcal{K} = \{1, \ldots, K\}$, and infinite capacity buffers are associated with each class of jobs. Class k jobs are served by a unique station $c(k)$ and their service times, $\{v_k(n), n \geq 1\}$, are positive, IID random variables drawn from some general distribution, with rate $\mu_k := 1/\mathbb{E}[v_k(1)] = 1/m_k$ and finite variance s_k. Upon completion of service at station $c(k)$, a class k job becomes a job of class m with probability P_{km} and exits the network with probability $1 - \sum_m P_{km}$, independent of all previous history. Assume that the general routing matrix $P = [P_{km}]$ is transient (that is, $I + P + P^2 + \ldots$ is convergent). Let $\{\phi^k(n)\}$ denote the sequence of K-dimensional IID Bernoulli random vectors such that $\phi^k_j(n) = 1$ if upon service completion the n^{th} class k job becomes a class j job and is zero otherwise, and let $\Phi^k(n) = \sum_{j=1}^n \phi^k(j)$. We set $\Phi^k(0) = 0$ and $\Phi = [\Phi^1, \ldots, \Phi^K]$. Every job class k can have its own renewal arrival process with interarrival times $\{u_k(n), n \geq 1\}$, with rate $\lambda_k := 1/\mathbb{E}[u_k(1)]$ and finite variance a_k. The set of classes that have a non-null exogenous arrival process will be denoted by $\mathcal{E}$. It is assumed that $\mathcal{E} \neq \emptyset$ and that the processes (v, u) are mutually independent.

For future reference, $\lambda = (\lambda_1, \ldots, \lambda_K)'$, where $'$ denotes a transpose, and $D = \mathbf{diag}\{m_1, \ldots, m_K\}$. C will be a $S \times K$ incidence matrix that maps classes into servers defined as $C_{ik} = 1$ if $c(k) = i$, and $C_{ik} = 0$, otherwise. The vectors of effective arrival rates and traffic intensities are given by

$$\alpha = (I - P')^{-1}\lambda \quad \text{and} \quad \rho = CR^{-1}\lambda < e,$$

where $R = (I - P')D^{-1}$, ρ_i denotes the nominal load (or utilization level) for server i, and e is the vector of ones of appropriate dimension. Hereafter, it will be assumed that $\alpha > 0$.

Denote by $Q_k(t)$ the total number of class k jobs in the system at time t, and by $Q(t)$ the corresponding K-vector of "queue lengths." A generic value for $Q(t)$ will be denoted by q. We define $|Q(t)| = \sum_k Q_k(t)$. The controller has discretion with respect to sequencing decisions. We will allow server splitting; that is, each server i can instantaneously divide its capacity into fractional allocations devoted into processing classes $k \in C_i$. Jobs within a class are served on First-In-First-Out (FIFO) basis and preemptive-resume type of service is assumed. Specifically, under these assumptions, a scheduling policy takes the form of the K-dimensional cumulative allocation process $\{T(t),\ t \geq 0;\ T(0) = 0\}$, where $T_k(t)$ is the cumulative time allocated into processing class k jobs up to time t. The cumulative allocation process should be non-decreasing and non-anticipating; the latter implies that current allocations depend only on information available up to time t. The cumulative idleness process is defined by $I(t) = et - CT(t)$, $I(0) = 0$ and $I(t)$ is non-decreasing.

The control problem is to choose a cumulative allocation process $\{T(t), t \geq 0\}$ to minimize

$$\mathbb{E}\left(\int_0^\infty e^{-\gamma t} c' Q(t) dt\right), \tag{2.1}$$

where $\gamma > 0$ is a positive discount factor and $c > 0$ is a vector of positive linear holding costs.

2.2. The associated fluid and Brownian model

Let $E_k(t)$ be the cumulative number of class k arrivals up to time t, $V_k(n)$ be the total service requirement for the first n class k jobs and define $S_k(t) = \max\{n : V_k(n) \leq t\}$ be the number of class k service completions when the appropriate server has allocated t time units in processing these jobs. The long run nominal allocation for class k jobs is given by $\alpha_k t/\mu_k$, or in vector form by $\alpha t = R^{-1}\lambda t$. That is, in order to cope with the external arrival processes the allocation control $T(t)$ should be approximately equal to αt. Let $\delta(t) = \alpha t - T(t)$ be the deviation control that measures how far is the actual allocation $T(t)$ from that vector of nominal time allocations. Then, the queue length dynamics are given by:

$$Q(t) \ = \ z + E(t) - S(T(t)) + \Phi(S(T(t))) \ = \ z + \chi(t) + R\delta(t), \tag{2.2}$$

where $\chi(t) = (E(t) - \lambda t) - \big[S(T(t)) - \Phi(S(T(t))) - (I - P')D^{-1}T(t)\big]$ and where z is the initial queue length vector. Note that when $\rho = e$, $I(t) = C\delta(t)$.

The fluid and Brownian models are obtained by approximating the centered term $\chi(t)$ by different quantities. Applying the functional Strong Law of Large Numbers (SLLN) for the processes (E, V, Φ) gives the associated fluid model. [2] Let q, y and $\bar{T}$ denote the fluid model variables that correspond to Q, δ and T of the stochastic network. The fluid model equations are given by

$$q(t) = z + Ry(t), \quad y(t) = \alpha t - \bar{T}(t), \quad \dot{\bar{T}}(t) \geq 0, \quad C\dot{\bar{T}}(t) \leq e. \tag{2.3}$$

Letting $\rho \to e$ and using the functional central limit theorem (CLT) for (E, V, Φ), we can approximate the centered term $\chi(t)$ by a Brownian motion, $X(t)$, with drift vector ν and covariance matrix

[2] The precise probabilistic assumptions required for the functional SLLN and CLT to hold and the scalings involved in deriving (2.3)-(2.4) will be spelled out in §5 where the main asymptotic optimality result is stated.

Σ. The problems that we will consider will admit Brownian optimal controls that do not depend on the drift and variance parameters, and for that reason we will not proceed here to characterize their values in terms of primitive problem data. Denoting by Z, Y, U the variables associated with Q, δ, I, respectively, system dynamics are approximated by

$$Z(t) = z + X(t) + RY(t), \quad Z(t) \geq 0, \quad U(t) = CY(t), \tag{2.4}$$

U is non-decreasing with $U(0) \geq 0$ and the control Y is adapted to the filtration generated by X.

The BCP is posed in terms of the equivalent workload formulation. The total workload for server i, denoted by $W_i(t)$, is defined as the expected total work for this server embodied in all jobs present in the system at time t, $W(\cdot) = MQ(\cdot)$, where $M = CR^{-1}$ is the workload profile matrix.

The equivalent workload formulation of the BCP is to choose RCLL (right continuous with left limits) processes Z and U to

$$\text{minimize } \mathbb{E}\left(\int_0^\infty e^{-\gamma t} c' Z(t) dt\right), \tag{2.5}$$

$$\text{subject to } W(t) = w_0 + \xi(t) + U(t), \tag{2.6}$$

$$W(t) = MZ(t), \quad Z(t) \geq 0, \tag{2.7}$$

$$U \text{ non-decreasing with } U(0) \geq 0, \tag{2.8}$$

$$U \text{ and } Z \text{ are adapted to } \xi, \tag{2.9}$$

where $w_0 = Mz$ and $\xi(t) = MX(t)$ is a $(M\nu, M\Sigma M')$ BM. The key insight that emerges from this formulation is that there is a dimension reduction in the control problem, the so called *state space collapse*, where the effective dimension drops from the number of classes, to the number of servers, that apart from simplifying the analysis, it directly offers powerful insights about the structure of good policies. The reader is referred to [19] for an introduction to Brownian approximating models, and to [24] for a discussion of equivalent workload formulations.

To summarize, we have described three models with different levels of detail specificity. The multiclass queueing model provides a "microscopic" system description, but leads to intractable problems in terms of control and/or performance analysis The deterministic fluid model that best captures the system transient response when the state is large. And finally, the Brownian model that approximates the steady state –or macroscopic– system behavior in heavy traffic. In terms of control, we will use the Brownian model to describe an asymptotically optimal policy, and the fluid model to translate the Brownian specification to an implementable tracking rule.

3. The class of "orthant" Brownian optimal controls

BCPs, while vastly simpler than the original problems at hand, are still quite intricate. Only a small subset of them admit closed form solutions. The numerical techniques due to Kushner and Dupuis [31] are a major breakthrough, but are still applicable to small problems. Promising recent results in

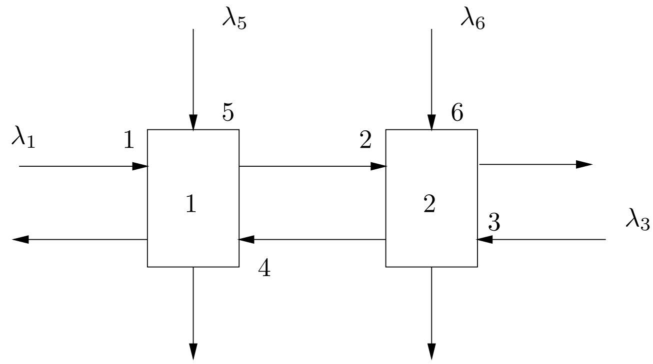

Figure 1. A two station example: $\lambda_k = .95$ and $\mu_k = 6$ for $k = 1, 3, 5, 6$, $\mu_2 = \mu_4 = 1.5$; $c = e$.

this domain have been reported by Kumar and Muthuraman [30]. Our focus is on network control problems whose associated BCPs, specified in (2.5)-(2.9), admit simple solutions. This restriction is made for two reasons: (a) while the Brownian solution is simple, the translation problem, which is our focus, is still subtle, and (b) the asymptotic analysis that will validate our translation proposal remains tractable. This is the natural first step in generalizing previous isolated examples that have been considered in the literature [25,26,21]. We first analyze the example of Figure 1 to highlight the relevant features of such controls, and then give a general definition for the class of problems we study in §§4-6. This restriction will be relaxed in §7.

A motivating example

We will start by analyzing the example shown in Figure 1. This is a variant of the network studied in [28,40] with 2 servers, 6 job classes, and where the flow of work is depicted by the corresponding arrows. The heavy traffic regime is approached when $\lambda_k = 1$ for $k = 1, 3, 5, 6$. Imitating the analysis given in [25, §3], we will show that the workload formulation (2.5)-(2.9) admits a surprisingly simple solution that is also *pathwise* optimal; i.e., it simultaneously minimizes $\sum_k Z_k(t)$ for all times t with probability one. In this example, the workload $W = MQ$ is given by

$$W_1 = m_1 Q_1 + m_4(Q_3 + Q_4) + m_5 Q_5 \quad \text{and} \quad W_2 = m_2(Q_1 + Q_2) + m_3 Q_3 + m_6 Q_6.$$

The optimal policy can be decomposed in two parts:

(i) select the idleness process U to control $W(t)$ through (2.6); and

(ii) at any point in time, pick the queue length vector $Z(t)$ that holds the current workload $W(t)$, i.e., $W(t) = MZ(t)$, and minimizes $\sum_k Z_k(t)$ through

$$Z(t) := \Delta(W(t)) = \operatorname{argmin} \left\{ \sum_k z_k : Mz = W(t),\ z \geq 0 \right\}. \tag{3.1}$$

That is, as W changes due to the random fluctuations in ξ and the control U, the controller instantaneously swaps Z's according to (3.1). From (2.4), instantaneous displacement between two vectors z_1, z_2 that hold the same workload (i.e., $M(z_1 - z_2) = 0$) can be achieved by applying the jump control $dY = R^{-1}(z_1 - z_2)$; this leaves the workload process unaffected since $dU = CdY = 0$.

The optimal Brownian control, which will be denoted by a superscript *, is given by:

$$U_i^* \text{ can increase only when } W_i^*(t) = 0 \quad \Rightarrow \quad \int_0^\infty W_i^*(t) dU_i^*(t) = 0, \text{ for } i \in \mathcal{I}, \tag{3.2}$$

which implies that

$$U_i^*(t) = \left[- \inf_{0 \le s \le t} w_{0,i} + \xi_i(s) \right]^+, \quad \text{for } i \in \mathcal{I}, \tag{3.3}$$

where $[x]^+ = \max(0, x)$; and given $W^*(t)$, set $Z^*(t) = \Delta(W^*(t))$, where Δ was defined in (3.1). The state space of the workload process is denoted by $\mathcal{S}^* = \mathbb{R}_+^2$. Conditions (3.2)-(3.3) imply that the controller uses the least amount of idleness to keep W^* in $\mathcal{S}^*$, idling only on its boundaries. For future use, J^* will denote the optimal cost $J^* \triangleq \mathbb{E}(\int_0^\infty e^{-\gamma t} c' Z^*(t) dt)$.

We note in passing that the mapping from W^* to the set of minimizers of the LP in (3.1) is continuous, and moreover, there exists a continuous selection function from W^* to Z^*; see Böhm [5, Theorem 2]. Hence, for our purposes Δ is a one-to-one continuous mapping.

The proof of optimality of (3.1)-(3.2) is as follows. First, we argue that U^* is feasible. Observe that any workload vector $w \ge 0$, is achievable through the queue length configuration $z = [0, 0, 0, 0, w_1/m_5, w_2/m_6]$. This implies there exists a feasible control U that satisfies (3.2). It is now well known that under the restriction that U is a RCLL process that satisfies (2.8), this control is unique and it is given by the reflection (or regulator) mapping in (3.3); see, for example, [18, §2.2]. In addition, $U_i^*(t)$ is minimal for all t with probability one, over all admissible controls that satisfy (2.6)-(2.9). Hence, $W_i^*(t) = w_{0_i} + \xi_i(t) - \inf_{0 \le s \le t} \xi_i(s)$, is non-negative, and it will also be the pointwise minimum over all $W(t)$ that satisfy (2.6)-(2.9) for all t with probability one.

To show that (U^*, Z^*) is indeed optimal we show that it is never cost-effective to intentionally increase workload. By the minimality of W^* and Z^*, this implies that (U^*, Z^*) is indeed optimal. Denote by W the workload vector at some time t, and consider the LP in (3.1):

$$\min \quad \{ \sum_k z_k : m_1 z_1 + m_4(z_3 + z_4) + m_5 z_5 = W_1, \quad m_2(z_1 + z_2) + m_3 z_3 + m_6 z_6 = W_2, \quad z \ge 0 \}.$$

Denoting dual variables by π, the dual LP is given by

$$\max \quad \{ W'\pi : m_1\pi_1 + m_2\pi_2 \le 1, \ m_2\pi_2 \le 1, \ m_4\pi_1 + m_3\pi_2 \le 1, \ m_4\pi_1 \le 1, \ m_5\pi_1 \le 1, \ m_6\pi_2 \le 1 \}.$$

The dual variable π_i is interpreted as the cost increase in $c'z$ per unit increase in w_i. By inspection we see that the solution of this dual LP is non-negative whenever $W \ge 0$ provided that

$$\max(m_2, m_6) > m_3 \quad \text{and} \quad \max(m_4, m_5) > m_1,$$

Note that the parameters in Figure 1 satisfy this condition. That is, $\min\{\sum_k z_k : Mz \ge w, z \ge 0\} = \min\{\sum_k z_k : Mz = w, z \ge 0\}$. Together with the minimality of U^* and W^* it follows that (U^*, Z^*) minimizes $\sum_k Z_k(t)$ over all policies for all t with probability one.

The general class of orthant optimal controls

This paper restricts attention to problems whose associated BCPs admit solutions of the form discussed above. These are the so called "orthant" optimal controls. In a nutshell, this requires that the (least) idleness control U^{min} is feasible and that the cost function is such that the system should never idle intentionally.

Definition 1. The BCP (2.5)-(2.9) has an "orthant" pathwise solution if $U^*(t) = U^{\min}(t)$, where

$$U_i^{\min}(t) = \left[-\inf_{0\leq s\leq t} w_{0_i} + \xi_i(s) \right]^+ \quad \text{for } i \in \mathcal{I}, \tag{3.4}$$

$W^*(t) = w_0 + \xi(t) + U^*(t)$, and

$$Z^*(t) = \Delta(W^*) \triangleq \text{argmin } \{c'z : Mz = W^*, \ z \geq 0\}. \tag{3.5}$$

In this case, (U^*, Z^*) minimizes total cost up to any time horizon with probability one. As a special case, (U^*, Z^*) also optimizes the infinite horizon discounted criterion in (2.5).

Proposition 3.1. The BCP (2.5)-(2.9) has an "orthant" pathwise solution if $U^{\min}(t)$ is feasible, and if for any workload vector $w \geq 0$, $\min\{c'z : Mz \geq w, \ z \geq 0\} = \min\{c'z : Mz = w, \ z \geq 0\}$. These conditions are equivalent to the following algebraic requirements:

$$\{w : w = Mz, \ z \geq 0\} = \mathbb{R}_+^S \qquad \text{and} \qquad \exists \, \pi \geq 0, \text{ such that } M'\pi = c. \tag{3.6}$$

The first part of (3.6) ensures that $U^{\min}$ is feasible (see also Yang [47]), and the second is a restatement of (3.6) using Farka's Lemma. Both conditions depend only on first moment information through the workload matrix M and the cost rate vector c, and can readily be checked using the original problem data. This is related to the definitions given in Meyn [38, §3]. This is, of course, a restricted set of problems that is even a subset of BCP's that admit *pathwise* optimal controls. To see this, consider a network control problem for a system with two servers in tandem, where $\mu_1 = \mu_2 = 1$, $\lambda = 1$, and unit holding costs (that is, $c'z = z_1 + z_2$). In this case, the non-idling policy is pathwise optimal; that is, it minimizes $Z(t)$ for all times t with probability one. However, it is not an orthant optimal, since the workload region $\mathcal{S}^* = \{w : w_2 = z_1 + z_2 \geq z_1 = w_1\}$ is the upper 45^o wedge and not the entire positive orthant, and thus $U^{\min}$ is not feasible.

4. Translation of "orthant" Brownian controls to tracking policies

We now restrict attention to multiclass queueing networks that satisfy the assumptions of §2.1, and whose associated BCPs satisfy the conditions of Definition 1, and consequently, admit orthant Brownian optimal solutions. Such controls require that: (i) server idleness is only incurred when total server workload is zero; and (ii) given $W^*(t)$, the system should instantaneously move to $Z^*(t)$ which is the minimum cost configuration for that workload position. Obviously, both of these statements are not meaningful in the context of the stochastic network when taken literally. For example, total workload for a server may be non-zero, while there are no jobs queued at the server for immediate service. Traditionally, the next step would be a problem-by-problem (often non-obvious) interpretation of the detailed solution of (3.5) that exploits problem structure and would not generalize outside the realm of the specific example considered.

Instead, the proposed translation is general, and is easy to implement and analyze. The basic structure of a continuous-review policy is simple: at any given time, the controller chooses a target, $z(q)$, using the solution of the Brownian control problem, and then selects via a fluid model analysis an allocation vector, $v(q)$, to steer the system in the appropriate direction. The control $v(q)$ stays constant between state transitions, whereat q changes and $z(q), v(q)$ are recomputed. This is akin to state-dependent sequencing rules, with the difference that rather than potentially switching priorities after state transitions, here we will be changing the fractional capacity allocations devoted into processing the various job classes. As an example one can consider a web-server processing requests from different classes of customers (e.g., repeat vs. new customers) that are associated with different processing requirements and potential revenues, and where at any point in time the system manager decides how much processing capacity to allocate to each customer class.

An important aspect of these policies that has not been exposed so far is the use of small safety stocks that prevent the queues from getting completely depleted. This is done in order to avoid the "boundary behavior" in the original system that is very different from that of its fluid idealization. The use of small safety stocks has been proposed in the past, first in the context of threshold policies in [27] and [4], and later on in discrete-review policies in [20,34] and [13,38].

The remainder of this section first provides some background material on fluid model minimum time control, which will be used to compute the allocation $v(q)$ given any desired target, and then gives a detailed description of the proposed policy.

4.1. Preliminaries: Fluid model minimum time control

We take a small detour to review *minimum time* control in fluid models. This material can be found in [34,13,38]. Specifically, we show that given an initial condition z and a feasible target state z^∞, the minimum time control is to move along a straight line from z to its target; this provides a very simple way of mapping a target state to an instantaneous control. Let t^* be the time it takes to reach z^∞ under some feasible control. That is, $q(t^*) = z^\infty = z + Ry(t^*)$, which implies that $y(t^*) = R^{-1}(z^\infty - z)$. From (2.3) it follows that $y(t^*) \le \alpha t^*$ which implies that

$$t^* \;\ge\; \max_k \frac{(R^{-1}(z^\infty - z))_k}{\alpha_k} \;\triangleq\; t^{(1)}(z, z^\infty). \tag{4.1}$$

The capacity constraint becomes that $Cy(t^*) = M(z^\infty - z) \ge (\rho - e)t^*$. This implies that

$$t^* \ge \max_{i:\rho_i \ne 1} \frac{(M(z^\infty - z))_i}{\rho_i - 1} \;\triangleq\; t^{(2)}(z, z^\infty), \tag{4.2}$$

and that for every i such that $\rho_i = 1$, $(M(z^\infty - z))_i \geq 0$; that is, the target workload for such servers can only increase. This follows from the fact that when $\rho_i = 1$, $w_i(t) = (Mq(t))_i = w_{0_i} + u_i(t)$, which is non-decreasing. When $\rho = e$, this constraint reduces to $Mz^\infty \geq Mz$.

The value $t^{\min}(z, z^\infty) = \max(t^{(1)}(z, z^\infty), t^{(2)}(z, z^\infty))$ is a lower bound for t^*. It is easy to verify that $t^{\min}$ is attained by moving along a straight line from z to z^∞. The corresponding trajectory and cumulative allocation process $\Psi^{\min}$ and $T^{\min}$, respectively, are given by

$$\Psi^{\min}(t; z, z^\infty) = z + (z^\infty - z)\frac{t \wedge t^{\min}(z, z^\infty)}{t^{\min}(z, z^\infty)}, \tag{4.3}$$

and

$$T^{\min}(t; z, z^\infty) = \alpha t - R^{-1}(z^\infty - z)\frac{t \wedge t^{\min}(z, z^\infty)}{t^{\min}(z, z^\infty)}. \tag{4.4}$$

4.2. Definition of the tracking policy

Safety stock requirements: We will try to maintain a minimum amount of work equal to θ at *each* buffer. Assuming that the interarrival and service time processes have exponential moments (this is made explicit in §5) it is sufficient that $\theta \sim \mathcal{O}(\log \mathbb{E}|Q|)$, where $\mathbb{E}|Q|$ is the expected amount of work in the system. Hence, θ is small compared to expected queue lengths in the system. With slight abuse of notation, we write $1-\rho$ for $1-\max_i \rho_i$. Classical queueing theory results suggest that $\mathbb{E}|Q| \sim 1/(1-\rho)$ and θ is set equal to

$$\theta := c_\theta \log(\frac{1}{1-\rho}), \tag{4.5}$$

where c_θ is a positive constant that depends on the probabilistic characteristics of the arrival and service time processes (see §5), and on the parameter values for λ and μ.

Planning horizon: Tracking policies will select a target of where the system should optimally be at some point in the future, and then make sequencing decisions to try to get there. The choice of target serves two purposes: a) enforce the safety stock requirement (i.e., avoid the boundaries); and b) optimize steady state performance. There is a tradeoff between short horizons that will tend to favor the former, over longer ones that focus on the latter. The planning horizon is set equal to

$$l := c_l \frac{\theta^2}{|\lambda|}, \tag{4.6}$$

for some constant $c_l > 0$, and where $|\lambda| = \sum_{k \in \mathcal{E}} \lambda_k$, serves as proxy for the "speed" of the network.

Finally, the policy is defined as follows. For all $t \geq 0$, we observe $Q(t)$, which will be denoted by q, and set $w = Mq$. The constraint $q \geq \theta$ implies that $w = Mq \geq \eta$, where $\eta \triangleq M\theta$. By enforcing safety stock requirements we effectively shift the region $\mathcal{S}^*$ to

$$\mathcal{S}^\theta = \{w + \eta : w \in \mathcal{S}^*\}. \tag{4.7}$$

Given w, the Brownian control would instantaneously move from q to $\Delta(w)$. Our policy, instead, chooses a control to keep $w \in \mathcal{S}^\theta$, $z(q) \geq \theta$, and track the target $\Delta(w)$. Let $q^\theta = [q - \theta]^+$.

1. *Idling control (reflection):* Given w, choose a target workload position $w^\dagger \in \mathcal{S}^\theta$ as follows

$$w \;\to\; w^\dagger = \;\max(w-(e-\rho)l,\eta) \;=\; \eta + w^\theta, \tag{4.8}$$

where $w^\theta = [w-(e-\rho)l-\eta]^+$. In heavy traffic, $w^\dagger = \max(w,\eta) = [w-\eta]^+ + \eta$.

2. *Target setting:* Translating the target workload $w^\dagger$ into the minimum cost queue length configuration is easy: the workload vector η is mapped to the vector θ of safety stocks, while w^θ is mapped to $\Delta(w^\theta) = \text{argmin}\left\{\sum_k \tilde{z}_k : M\tilde{z} = w^\theta, \tilde{z} \geq 0\right\}$. The target state is set to

$$z(q) \;=\; \theta + \Psi^{\min}(l;q^\theta,\Delta(w^\theta)) \;=\; \theta + \underbrace{q^\theta + \left(\Delta(w^\theta) - q^\theta\right)\frac{l \wedge t^{\min}(q^\theta,\Delta(w^\theta))}{t^{\min}(q^\theta,\Delta(w^\theta))}}_{\text{position at time } t+l \text{ on min-time path } q^\theta \to \Delta(w^\theta)}. \tag{4.9}$$

3. *Tracking:* Given the target $z(q)$, choose a vector instantaneous fractional allocations in order to steer the state from q to $z(q)$ in minimum time as follows

$$v(q) = \frac{T^{\min}(l;q,z(q))}{l}. \tag{4.10}$$

Note that $CT^{\min}(l;\cdot,\cdot) \leq le$, and thus $Cv(q) \leq e$ satisfies the capacity constraints. The actual effort allocated to class k will be $\dot{T}_k(t) = v_k(q)\cdot \mathbf{1}_{\{q_k>0\}}$, where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. That is, unused capacity is not redistributed to other classes. This gives the system the capability of intentionally idling a server when that seems beneficial.

Summing up, a tracking policy is uniquely defined through equations (4.5)-(4.10), and the system parameters λ and μ. The target is computed through (4.9), and the corresponding instantaneous control using (4.10). Note that the target and control $z(q), v(q)$ only change when the queue length vector changes. Hence, the control only needs to be computed upon state transitions. Also, following the remark made after (3.1), there exists a continuous selection function that maps $w^\dagger$ to z^*. Using (4.4), we conclude that the tracking policy that maps q into $v(q)$ is a measurable and continuous function. It is also a *head-of-line* policy; that is, only the first job of each class can receive service at any given time.

We conclude with a few comments regarding the use of safety stocks, the planning horizon and the choice of target workload, $w^\dagger$ that encompasses the idling control of this policy. First, note that the safety stock requirement is a "soft" constraint and the queue length vector may indeed drop below θ. Equations (4.8)-(4.10) are still valid when $q \not\geq \theta$, and at such times the system manager will adjust its sequencing decisions to bring the state close to θ again. Second in terms of the planning horizon, other choices for l would also work. The key is for l to be between the time it takes to accumulate the safety stock (of order $\theta/|\lambda|$) and the time required to move from one queue length position to another that is longer, of order $\mathbb{E}|Q|/|\lambda|$. Finally, note that the fluid model dynamics can be rewritten in the form

$$w(t) = Mq(t) = w(0) + u(t) + (\rho - e)t. \tag{4.11}$$

First, we focus on the case $\rho = e$. Staring from w, the controller chooses the target workload position $w^\dagger \in \mathcal{S}^\theta$, then selects a target queue length $z(q)$ to hold this workload, and a control $v(q)$ to get there.

By construction, $z(q)$ hold the same workload as $w^\dagger$, and from (4.11), this implies no idling. Hence, $I(t)$ will only increase if (i) $w < \eta$, in which case $[\eta_i - w_i]^+$ time units of idleness are "forced" into each server i (this corresponds to vertical displacement along the U_i^* direction of control that was described in (3.2)) or (ii) when the controller plans to allocate $v_k(q) > 0$ into processing class k jobs, but $q_k = 0$. Case (ii) corresponds to "unplanned" idleness, where the policy would have liked to work on a job class which, however, happened to be empty at this time. The safety stock requirement will make (ii) negligible, and hence, the servers will only idle when w is "reflected" back to $w^\dagger$ on the boundary $\partial\mathcal{S}^\theta$. This is precisely what the Brownian control policy prescribed, only shifted by η, and corresponds to idling a resource when there is no work for that resource anywhere in the network. At such times the policy knows that the resource must idle, and incorporates that into its allocation decision. Finally, if $\rho < e$, keeping workload constant -as it is prescribed from the Brownian policy- will result into idling of the servers at a rate of $e - \rho$. To correct for this effect, we try to set $w^\dagger$ through (4.8) that tries to reduce the workload by using this excess capacity.

5. Heavy traffic asymptotic optimality of tracking policies

This section presents the main result of this paper, namely that the proposed continuous-review policy is asymptotically optimal as the system approaches the heavy traffic regime. This first involves the introduction of some new notation and the definition of a sequence of systems that approach heavy traffic in an appropriate sense.

5.1. Heavy-traffic scaling and the Brownian approximating model

The network will approach the heavy traffic operating regime as

$$\lambda \to \infty, \quad \mu \to \infty \quad \text{and} \quad \rho \to e,$$

This type of heavy traffic scaling can arise from a "speed up" of the network operation, which could be the outcome of a technological innovation, or simply a subdivision of individual jobs in finer tasks (packets) that arrive and get processed faster.

We shall consider a sequence of systems that correspond to different values of (λ, μ) that will be indexed by $|\lambda|$, which, as mentioned earlier, serves as proxy for the "speed" or "size" of the network. To simplify notation we will denote $|\lambda|$ by r, and attach a superscript "r" to quantities that correspond to the r^{th} system; i.e., the sequences of random variables are now $\{u_l^r(i)\}$ for all $l \in \mathcal{E}$ and $\{v_k^r(i)\}$ for all $k \in \mathcal{K}$, the arrival and service rate vectors will be λ^r and μ^r, and the vector of variances will be denoted by a^r and s^r. The heavy traffic limit is approached as

$$r := |\lambda^r| \to \infty, \quad \frac{\lambda^r}{r} \to \lambda \geq 0, \quad \frac{\mu^r}{r} \to \mu > 0 \quad \text{and} \quad \sqrt{r}(\rho^r - e) \to \nu, \tag{5.1}$$

where ν is a finite S-dimensional vector. For the example of Figure 1, $\lambda_k = 1$ for $k = 1, 3, 5, 6$ and μ is equal to the rates given in the figure caption; this makes the traffic intensity equal to 1 at both stations. The speed of convergence of $\rho^r \to e$ is what typically appears in heavy traffic analysis and leads to the appropriate Brownian limit. Similarly, $R^r = (I - P')(D^r)^{-1}$ and $M^r = C(R^r)^{-1}$ and their corresponding limits will be $\frac{1}{r}R^r \to R := (I - P')D^{-1}$ and $rM^r \to M := CR^{-1}$.

For the functional strong law and central limit theorems to hold for the appropriately scaled processes, we will impose the following conditions taken from [7,45]:

$$\forall\ l \in \mathcal{E},\quad r^2 a_l^r \to a_l < \infty \quad \text{and} \quad \forall\ k \in \mathcal{K},\quad r^2 s_k^r \to s_k < \infty,$$

and

$$\sup_{l\in\mathcal{E}} \sup_{r} \mathbb{E}\left[r^2 u_k^r(1)^2; u_k^r(1) > \frac{n}{r}\right] \to 0 \quad \text{and} \quad \sup_{k\in\mathcal{K}} \sup_{r} \mathbb{E}\left[r^2 v_k^r(1)^2; v_k^r(1) > \frac{n}{r}\right] \to 0 \text{ as } n \to \infty;$$

the latter are Lindeberg type conditions.[3] For the r^{th} set of parameters, let $V_k^r(n)$ be the total service requirement for the first n class k jobs for $k \in \mathcal{K}$, $V_k^r(t) = V_k^r(\lfloor t \rfloor)$, and $E_l^r(t)$ be the number of class l arrivals in the interval $[0, t]$ for $l \in \mathcal{E}$. We require that uniform large deviations upper bounds hold for the sample paths of the processes $E_l^r(\cdot)$ and $V_k^r(\cdot)$. Specifically, for any $l \in \mathcal{E}$ or $k \in \mathcal{K}$, $\epsilon > 0$ and t sufficiently large, there exists $r^* > 0$ such that for all $r \geq r^*$,

$$\mathbb{P}\left(\sup_{0\leq s\leq t/r} |E_l^r(s) - \lambda_l^r s| > t/r\epsilon\right) \leq e^{-f_a^k(\epsilon)t/r} \tag{5.2}$$

$$\mathbb{P}\left(\sup_{0\leq s\leq t/r} |V_k^r(s) - m_k^r s| > t/r\epsilon\right) \leq e^{-f_s^k(\epsilon)t/r}, \tag{5.3}$$

where $f_a^k(\epsilon), f_s^k(\epsilon) > 0$ are the large deviations exponents. Similar bounds are known to hold for IID Bernouli routing random variables; the corresponding exponents are denoted by f_r^{kl}. [4]

We define the scaled processes by

$$Z^r = \frac{Q^r}{\sqrt{r}}, \quad U^r = \sqrt{r}I^r, \quad Y^r = \sqrt{r}\delta^r \quad \text{and} \quad W^r = \sqrt{r}M^r Q^r = rM^r Z^r, \tag{5.4}$$

consider initial conditions of the form $Q^r(0) = r^{1/2}z^r$ for some $z^r \geq 0$, and study the limits as $r \to \infty$ of the network processes Z^r, U^r, Y^r. Our analysis will also assume the technical condition that

[3] Both [7,45], allow the distributions of the first (residual) interarrival time and the residual processing times for jobs present at time $t = 0$ to be different than those of u^r and v^r, respectively. For simplicity we assume that residual service times of all class k jobs present at $t = 0$ are drawn from the sequence $\{v_k^r\}$ for $k \in \mathcal{K}$, and that the first class l arrival occurs at time $t = u_l^r(1)$ for $l \in \mathcal{E}$. In this case, conditions [45, (80)-(81)] and [7, (3.5)] are automatically met.

[4] For fixed r and i.i.d. sequences of random variables with exponential moments, Mogulskii's Theorem [16, Thm 5.1.2] implies (5.2)-(5.3); see [15] for extensions to weakly dependent sequences. The uniform bound guarantees that the families $\{u_l^r\}$ and $\{v_k^r\}$ have a guaranteed rate of convergence for large r. The exponents in (5.2)-(5.3) are taken as primitive data that describe the burstiness of the interarrival and service time processes, as is done in communication systems.

$\lim_{r\to\infty} |Z^r(0) - \Delta(W^r(0))| \to 0$ in probability. Note that (5.4) only involves scaling in the spatial coordinate and not in time.[5]

5.2. Main result: heavy traffic asymptotic optimality

The heavy traffic regime is approached as the arrival rates and the processing capacities grow according to (5.1). Recall that a policy π can be described by its allocation process $\{T(t), t \geq 0\}$. Hence, $\pi^r = \{T^r(t), t \geq 0\}$ will denote the policy when the aggregate arrival rate is $|\lambda| = r$ and the service rate vector is μ^r, and $\{\pi^r\}$ will refer collectively to the corresponding sequence of policies –one for each set of parameters. The sequence of tracking policies will be denoted by $\pi^{r,*}$. In the sequel, the expectation operator is taken with respect to the probability distribution in the path space of queue lengths induced by the policy π; this will be denoted by $\mathbb{E}_\pi$.

Theorem 1. Consider multiclass networks that satisfy the assumptions in §2.1, §5.1, and the conditions of Definition 1. Let $\{\pi^r\}$ denote any sequence of control policies. Then,

$$\liminf_{r\to\infty} \mathbb{E}_{\pi^r} \left(\int_0^\infty e^{-\gamma t} \sum_k c_k Z_k^r(t) dt \right) \geq \lim_{r\to\infty} \mathbb{E}_{\pi^{r,*}} \left(\int_0^\infty e^{-\gamma t} \sum_k c_k Z_k^{r,*}(t) dt \right) = J^*. \tag{5.5}$$

That is, the tracking policy of (4.5)-(4.10) is asymptotically optimal in the heavy traffic limit.

This is the main analytical result of the paper. It asserts that as the model parameters approach a limiting regime where the network operates in heavy traffic and the model approximation invoked in §2 becomes valid, the performance under the proposed policy (that also changes as a function of the model parameters) approaches that of the optimally controlled Brownian model that one started with. In this sense, this asymptotic performance certificate validates the proposed policy as a translation mechanism for the Brownian control policy to the original network. This is one of the first such results that is proved for a general translation mechanism and applies to a wide range of problems, rather than just a specific example. In parallel work to ours, Meyn [38, Thm. 4.3-4.5] established a related result for a family of discrete-review policies. His result seems to hold for a more general class of network models, but also under the assumption that the associated fluid and Brownian formulations admit pathwise optimal solutions. The translation mechanism used in [38] and the asymptotic analysis technique are different to ours. All proofs are relegated to the appendix.

[5] This should be contrasted with the more traditional diffusion scaling where one considers system parameters that vary like $\lambda^r \to \lambda, \mu^r \to \mu$ such that $\rho \to e$, and uses the CLT type of scaling $\tilde{Z}^r(t) = Q^r(rt)/\sqrt{r}, \ldots$ (*). The two setups are completely equivalent. For example, to recover the latter starting from (5.1) and (5.4) we need to stretch out all "speeded up" interarrival and service time random variables by r and then use (*).

6. Discussion

We now give several remarks about the various features of the proposed family of continuous-review policies and their relation to policies derived via a fluid model analysis.

Scaling relations

From (5.4), we can deduce that

$$Q^r = \mathcal{O}(\sqrt{r}) \quad \text{and} \quad \theta^r = \mathcal{O}(\log r).$$

From $\sqrt{r}(\rho - e) \to \nu$, one gets that r is of order $\mathcal{O}(1/\sqrt{1-\rho})$, and as a result the above relations can be rewritten in the form $Q^r = \mathcal{O}(1/\sqrt{1-\rho})$ and $\theta^r = \mathcal{O}(\log(1/(1-\rho))$. That is, for systems operating close to heavy traffic, safety stocks become negligible compared to queue lengths.

In terms of the magnitude of the planning horizon, the time required to accumulate the safety stock is roughly of order $\log r/r$, while the time required to move from one queue length to another target position is of order $\sqrt{r}/r = 1/\sqrt{r}$. The choice for the planning horizon $l^r = \mathcal{O}(\log^2 r/r)$ lies between these two values. This illustrates the tradeoff eluded to in §4.2 between keeping the planning horizon short and effectively focusing on satisfying the safety stock requirement, and making it long and focusing on tracking the minimum cost state. The proposed choice is long such that the effect of tracking the minimum cost state dominates, but sufficiently short such that the safety stock requirement is met with high probability and undesirable "boundary behavior" of the network is negligible. This "separation of time scales" between the planning horizon and the time required to satisfy the safety stock requirement or track the target state differs from what is used in discrete-review policies [20,34] or more recently in [38]. There the length of each discrete-review period is of order $\log r/r$, and the idea is that processing plans within each period are fulfilled from work that is present at the beginning of each period. In contrast, continuous-review periods are less rigid about the enforcement of the safety stocks and are likely to do better in terms of cost optimization. Indeed, experimental evidence has shown that continuous-review implementations tend to perform better than their discrete-review cousins; see Harrison [21, §7]. Finally, this rule-of-thumb should apply in other application domains where such continuous-review policies are applicable.

Fluid scale behavior

The relations given above suggest that fluid scaled processes are defined by

$$q^r(\cdot) = \frac{Q^r(\cdot/\sqrt{r})}{\sqrt{r}}, \quad u^r(\cdot) = \sqrt{r} I^r(\cdot/\sqrt{r}), \quad y^r(\cdot) = \sqrt{r}\delta^r(\cdot/\sqrt{r}) \quad \text{and} \quad w^r(\cdot) = rM^r q^r(\cdot), \tag{6.1}$$

and that their respective limits should satisfy the fluid model equations together with some extra conditions that depend on the policy. Proposition 7.2 shows that this is indeed the case, and, in fact, that the fluid model trajectory matches the minimum time trajectory $\Psi^{\min}(\cdot; z, \Delta(w_0))$. This implies that in fluid scale the system attains its target $\Delta(w_0)$ in minimum time. Since the diffusion scaling in (5.4) does not involve any stretching of time as in (6.1), translation in finite time in fluid scale appears instantaneous in the diffusion model – the essence of the state space collapse property.

Minimum cost vs. minimum time tracking

So far we have used the solution of the Brownian control problem to select the minimum cost state $\Delta(w)$, and the minimum time control to compute the target $z(q)$ and to translate it to an allocation vector. The latter does not involve any cost considerations. From the viewpoint of heavy traffic asymptotic analysis, what is important is to track the appropriate target in finite time in fluid scale, so that it appears instantaneous in diffusion scaling. And, *any* fluid trajectory that reaches its target in finite time will do. In the original system, however, one may still benefit by following a minimum cost trajectory to compute the target $z(q)$ and the associated allocation.

The latter is computed through a fluid optimization procedure (see [34]) described below. For simplicity, we assume that the system is in heavy traffic; i.e., $\rho = e$, $w^\dagger = \max(w, \eta)$ and $w^\theta = [w-\eta]^+$. Starting from an initial condition z and given a feasible target state z^∞, the minimum cost trajectory is computed as the solution of

$$\bar{V}(z, z^\infty) = \min_{\bar{T}(\cdot)} \left\{ \int_0^\infty [c'q(t) - c'z^\infty]\, dt : q(0) = z \text{ and } (q, \bar{T}) \text{ satisfies (2.3)} \right\}. \tag{6.2}$$

$\bar{V}(z, z^\infty)$ is the minimum achievable cost, while $\Psi^*(\cdot; z, z^\infty)$ and $T^*(\cdot; z, z^\infty)$ will denote the minimum cost trajectory and the corresponding cumulative allocation process. Basic facts of optimization theory imply that the solution exists, it is Lipschitz continuous in (z, z^∞), and its time derivative $\dot{\Psi}^*(0; z, z^\infty)$ is continuous for almost all (z, z^∞); at points of discontinuity we assign $\dot{\Psi}^*(0; z, z^*) = \dot{\Psi}^*(0^+; z, z^*)$. (See [34, §6] for details and examples of optimal fluid control policies.)

A continuous-review policy based on minimum cost tracking is defined as follows:

(a) Solve the BCP (2.5)-(2.9). Given the queue length vector q, use the optimal Brownian control policy to choose the target queue length configuration $\theta + \Delta(w^\theta)$.

(b) Given $\Delta(w^\theta)$, choose the target state $z(q)$ at time l and the allocation $v(q)$ using the minimum cost trajectory from q^θ to $\Delta(w^\theta)$ that is computed from (6.2).

Step (a) is about "long term" (or steady state) behavior, while step (b) is about "short term" (or transient) response. The corresponding tracking policy is defined using the solution of the minimum cost tracking problem of (6.2), Ψ^* and T^*, in place of $\Psi^{\min}$ and $T^{\min}$.

Theorem 2. Consider multiclass networks that satisfy the assumptions in §2.1, §5.1 and the conditions of Definition 1. Let Ψ^* and T^* be the optimal trajectory and associated control for (6.2). Then, the tracking policy defined by (4.5)-(4.10) and using Ψ^* and T^* in place of $\Psi^{\min}$ and $T^{\min}$ is asymptotically optimal in the heavy traffic limit.

This policy using the minimum cost fluid control to compute targets and allocations, coincides with the policy that one would have derived by fluid model analysis alone. This is easy to verify by convincing oneself that for any initial condition the fluid model will also eventually steer the state to the minimum cost position $\Delta(w_0)$; i.e., it would track the same target along the minimum cost trajectory.

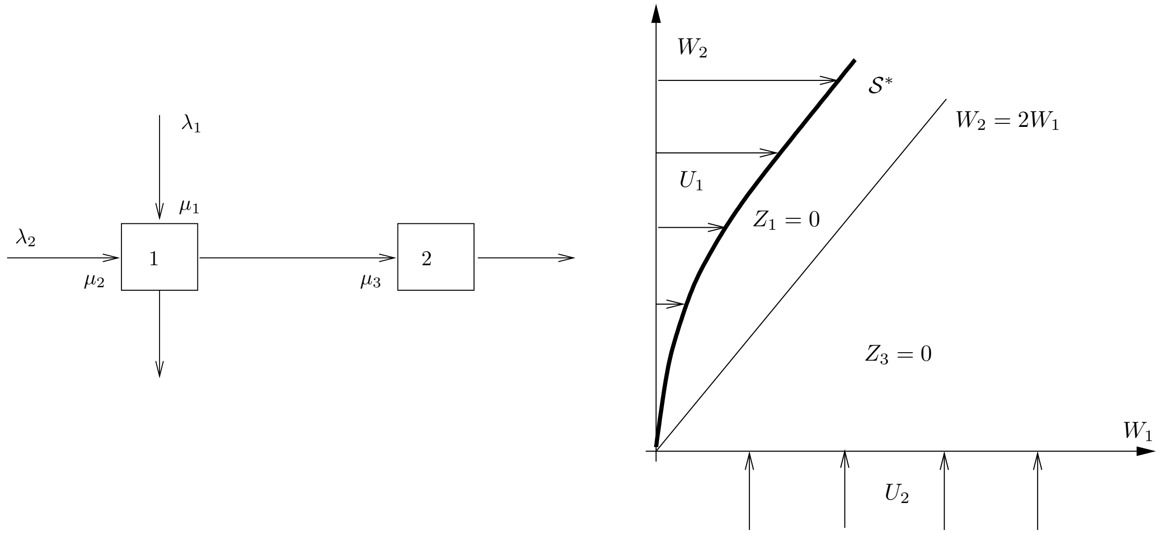

Figure 2. Left panel: *Criss-Cross* network with parameters $\lambda_1 = \lambda_2 = \rho,\ \mu_1 = \mu_2 = 2,\ \mu_3 = 1$. Right panel: Solution to Brownian control problem with the objective shown in (7.1).

Hence, Theorem 2 shows that fluid model analysis is equivalent to Brownian analysis for problems that admit "orthant" optimal controls. A related result was derived in Meyn [38].

7. Extension: Policy translation for problems that do not admit "orthant" solutions

The last point we address is the policy translation problem in cases where the Brownian control policy does not have the simple structure of an orthant control. To contrast with Definition 1: (a) the region $\mathcal{S}^*$ is no longer the positive orthant $\mathbb{R}_+^S$, but it is some general convex subset of $\mathbb{R}_+^S$; (b) the directions of control (or the "angles of reflection") when the workload process is on the boundary $\partial\mathcal{S}^*$ can be any linear combination of the idleness controls (U_i vectors); and (c) one has to specify what the control action should be if the initial condition is outside $\mathcal{S}^*$.

We start with a brief discussion of the so called criss-cross network shown in Figure 2 that was first introduced by Harrison and Wein [25]. It is a multiclass network with two servers and three job classes. Heavy traffic is approached as $\rho \to 1$. We focus on the objective

$$\mathbb{E}\left(\int_0^\infty e^{-\gamma t}(Q_1(t) + Q_2(t) + 2Q_3(t))dt\right). \tag{7.1}$$

In this case, condition (3.6) is no longer satisfied. Following the classification in [35, p.2136], the Brownian optimal solution is not an orthant control as in §3, but it will have a switching surface like the one depicted in Figure 2; this has been verified numerically by Kumar and Muthuraman [30]. The optimal Brownian control policy is the following: (a) keep $W(t) \in \mathcal{S}^*$; (b) server 2 incurs idleness only when $W_2(t) = 0$; (c) server 1 incurs idleness only when $W_1(t)$ hits the left boundary of $\mathcal{S}^*$; (d) if $w_0 \notin \mathcal{S}^*$ (i.e., server 1 initial workload lies above the boundary of $\mathcal{S}^*$), then server 1 incurs an instantaneous impulsive control $U_1(0) > 0$ such that $W_1(0)$ is shifted horizontally to $\partial\mathcal{S}^*$; (e) in the interior of $\mathcal{S}^*$, W behaves like the Brownian motion ξ; and (f) $Z = \Delta(W)$, where Δ is defined in (3.5) and $c'z = z_1 + z_2 + 2z_3$. The intuitive explanation of this policy was given in [30]. First, they note that the solution of the LP in (3.5) requires that $z_1^* = 0$ whenever the workload position falls above the line $w_2 = 2w_1$. Next consider the case $w_2 \gg w_1$. In this case, $Z_1 = 0$ and $Z_3 \gg Z_2$. Given that it is more expensive to hold jobs in queue 3 rather than queue 2, server 1 is better off just working on class 1 (keeping $Z_1 = 0$) and utilizing only half of its capacity, until the backlog in queue 3 has sufficiently decreased. This will involve some idling of server 1, thus pushing the server 1 workload process horizontally towards the boundary of region $\mathcal{S}^*$. When Z_3 falls below a certain level, server 1 will start serving both classes 1 and 2 and will not incur any further idleness.

To translate such a policy, we need to adjust the "reflection" step, where the controller chooses a target workload vector $w^\dagger$ based on the current workload w, such that: we keep $w^\dagger$ in the region $\mathcal{S}^*$; $w^\dagger$ increases along the appropriate directions of control when w is close to the boundaries of $\mathcal{S}^*$; and, $w^\dagger$ is selected appropriately when $w \notin \mathcal{S}^*$.

In general, we will provide a translation mechanism for Brownian control policies that are specified in the form of a Skorokhod problem as discussed by Dupuis and Ramanan [17, §2.2]:

(a) $W(t) = w_0 + \xi(t) + U(t)$.

(b) $\forall\, t \geq 0,\ W(t) \in \mathcal{S}^* \subseteq \mathbb{R}_+^S$.

(c) $|U|(t) = \sum_i U_i(t) < \infty$.

(d) $|U|(t) = |U|(0) + \int_{(0,t]} \mathbf{1}_{\{W(s)\in\partial\mathcal{S}^*\}} d|U|(s)$.

(e) For some measurable mapping $\nu : \mathbb{R}_+^S \to \partial\mathcal{S}^*$, for all $w_0 \notin \mathcal{S}^*$,

$$W(0) = \nu(w_0) \ \text{ and } \ U(0) = W(0) - w_0 = \nu(w_0) - w_0 \geq 0,$$

and for all $w_0 \in \mathcal{S}^*$, $\nu(w_0) = w_0$. Note that if $\mathcal{S}^* = \mathbb{R}_+^S$, then $\nu(w) = w$ for all $w \in \mathbb{R}_+^S$.

(f) Let $r(x)$ be the directions of allowed control when $W(t) = x \in \partial\mathcal{S}^*$ and $d(x) = \{y \in r(x) :\ \|y\| = 1\}$; if $x \notin \partial\mathcal{S}^*$, then $r(x) = \emptyset$ and $d(x) = \emptyset$. For some measurable function $\zeta : [0,\infty) \to \mathbb{R}_+^S$ such that $\zeta(t) \in d(W(t))$,

$$U(t) = U(0) + \int_{(0,t]} \zeta(s) dU(s).$$

(g) $Z = \Delta(W)$, where Δ was defined in (3.5).

Finally, for continuity we require that along any sequence of initial conditions $\{w_0^n\}$, $w_0^n \notin \mathcal{S}^*$, such that $\lim_n w_0^n = w^* \in \partial\mathcal{S}^*$, $\lim_n(\nu(w_0^n) - w_0^n)/\|\nu(w_0^n) - w_0^n\| \in r(\lim_n w_0^n) = r(w^*)$.

A Brownian control policy is specified through the mappings ν and ζ that give the directions of control when the workload position is outside $\mathcal{S}^*$, or on the boundary $\partial\mathcal{S}^*$, respectively. For the criss-cross example, $\mathcal{S}^*$ is the convex region in the interior of the control surface shown in "bold" print. Let $\delta(w_2) = \max\{w_1 \geq 0 : [w_1, w_2] \in \partial\mathcal{S}^*\}$; that is, the pairs $(\delta(w_2), w_2)$ define the boundary of $\mathcal{S}^*$. Then, if $w \notin \mathcal{S}^*$, $\nu(w) = [\delta(w_2), w_2]$, which corresponds to instantaneous displacement in a horizontal direction by idling server 1. The control directions on $\partial\mathcal{S}^*$ are: $d(w) = \{[1,0] : w_1 = \delta(w_2),\ w_2 > 0\}$ (i.e., idle server 1), $d(w) = \{[0,1] : w_1 > 0,\ w_2 = 0\}$ (i.e., idle server 2), and $d(w) = \{[1,1] : w_1 = w_2 = 0\}$ (i.e., idle both servers). Finally, d uniquely defines r and ζ.

Tracking policy implementation. Once again, $\eta = M\theta$, q denotes a generic value of the queue length vector, $w = Mq$, $q^\theta = [q-\theta]^+$, $w^\theta = [w-(e-\rho)l-\eta]^+$, and $\mathcal{S}^\theta = \{w+\eta : w \in \mathcal{S}^*\}$. A Brownian control policy specified through points (a)-(g) given above is translated as follows:

(i) *Reflection:* $w^\dagger = \eta + \nu(w^\theta)$.

(ii) *Target:* $z^* = \theta + \Delta(\nu(w^\theta))$ and $z(q)$ is set according to (4.9).

(iii) *Tracking:* $v(q)$ is set according to (4.10). (One can use either Ψ^*, T^* or $\Psi^{\min}, T^{\min}$ for (ii)-(iii).)

While a heavy traffic asymptotic optimality result has been proved for the “bad” parameter case of the criss-cross network in [32], the structure of their policy was problem specific and quite hard to implement. The development of a rigorous asymptotic analysis for this family of continuous-review policies and the establishment of an asymptotic optimality result for a more general class of network control problems (in the spirit of the results of Bramson and Williams and the analysis of Appendices A and B) is an interesting and challenging open problem.

Finally, we refer the reader to Chen *et.al* [11] for some interesting results that contrast the performance of the optimal Brownian control (when this takes the form shown in Figure 2) with the corresponding fluid optimal policy.

Acknowledgements: I am indebted to Mike Harrison for providing the initial motivation for this work and for many helpful comments and suggestions all along the course of this project. I am also grateful to Ruth Williams, Maury Bramson and Sunil Kumar for many suggestions and useful discussions, and to Mor Armony for pointing out reference [5].

Appendix A: Proofs

A few words on notation. Let $\mathbf{D}^d[0,\infty)$ with d any positive integer be the set of functions $x : [0,\infty) \to \mathbb{R}^d$ that are right continuous on $[0,\infty)$ with left limits on $(0,\infty)$, and $\mathbf{C}^d[0,\infty)$ denote the set of continuous functions $x : [0,\infty) \to \mathbb{R}^d$. For example, E, V are elements of $\mathbf{D}^K[0,\infty)$. For a sequence $\{x^n\} \in \mathbf{D}^d[0,\infty)$ and an element $x \in \mathbf{C}^d[0,\infty)$, we say that x^n converges uniformly on compact intervals to x as $n \to \infty$, denoted by $x^n \to x$ *u.o.c.*, if for each $t \geq 0$, $\sup_{0\leq s\leq t} |x^n(s) - x(s)| \to 0$ as $n \to \infty$. For two sequences $x^n,\ y^n$ of right continuous functions with left limits, we will say that $x^n(\cdot) \approx y^n(\cdot)$, if $|x^n(\cdot) - y^n(\cdot)| \to 0$, uniformly on compact sets, as $n \to \infty$. For a sequence of stochastic processes $\{X^r, r > 0\}$ taking values on $\mathbf{D}^d[0,\infty)$, we use $X^r \Rightarrow X$ to denote convergence of X^r to X in distribution.

We first provide a skeleton of the proof by reviewing the series of propositions that will need to be established and then given detailed proofs. The first step is a large deviations analysis under the proposed policy. The planning horizon for the r^{th} set of parameters is given by $l^r = c_l(\theta^r)^2/r$.

Proposition 7.1. (Large Deviations) For some appropriate constant $c_\theta > 0$ and r sufficiently large,

$$\mathbb{P}\left(\inf_{\kappa^r l^r \le s \le t} Q^{r,*}(s) \not\ge e\right) \le \frac{t}{r^2}, \tag{7.2}$$

where $\kappa^r > 0$ and $\kappa^r l^r \to 0$ as $r \to \infty$.

Broadly speaking, starting from any initial condition (maybe empty), the system will grow its target safety level position within $\kappa^r l^r$ time and then stay around that level for ever after with very high probability. This implies that safety stock levels of order $\mathcal{O}(\log(|\lambda|))$ are sufficient in the sense that as $|\lambda|$ increases, the probability that any of the queue lengths will get depleted becomes negligible. A consequence of this result is that asymptotically the system incurs no idleness by not being able to implement the control $v(q)$ due to empty buffers.

Following Bramson's analysis, the next step is to analyze the fluid model response of our proposed policy. We first show that in fluid scale, the system is tracking the minimum time trajectory $\Psi^{\min}$ from z to the target $\Delta(w_0)$, where $w_0 = Mz$ is the initial workload position. A modification of Bramson's arguments will then show that under diffusion scaling the queue length process is moving instantaneously to its target $\Delta(w)$.

Proposition 7.2. (Fluid scale analysis) Consider any sequence of initial conditions $\{r^{1/2}z^r\}$. For every converging subsequence $\{z^{r_j}\}$, where $q^{r_j}(0) = z^{r_j} \to z$ for some $z \ge 0$, and almost all sample paths

$$(q^{r_j,*}(\cdot), y^{r_j,*}(\cdot), u^{r_j,*}(\cdot)) \to (q(\cdot), y(\cdot), u(\cdot)) \quad \text{u.o.c..}$$

Moreover, $q(0) = z$ and $q(\cdot) = \Psi^{\min}(\cdot; z, \Delta(w_0))$, where $w_0 = Mz$.

Proposition 7.3. (State space collapse) Assume that $|Z^r(0) - \Delta(W^r(0))| \to 0$ in probability as $r \to \infty$. For all $t \ge 0$,

$$\sup_{0 \le s \le t} |Z^{r,*}(\cdot) - \Delta(W^{r,*}(\cdot))| \to 0, \quad \text{in prob. as } r \to \infty. \tag{7.3}$$

To establish the diffusion limit we need to show that the system incurs non-negligible amounts of idleness only when $W^{r,*}$ is close to the boundary of $\mathcal{S}^{\theta^r/\sqrt{r}}$. In the heavy traffic regime, the difference between $\mathcal{S}^{\theta^r/\sqrt{r}}$ and $\mathcal{S}^*$ becomes negligible, and the pair $(W^{r,*}, U^{r,*})$ will asymptotically satisfy (3.2). This determines the limiting workload and idleness processes, and together with state space collapse, establishes the desired result. We proceed as follows.

Recall that $W^r = \sqrt{r}M^r Q^r = rM^r Z^r$. Using (2.2) and (5.4) we get that

$$W^r(t) = W^r(0) + rM^r\left(\frac{E(t) - \lambda t}{\sqrt{r}} - \frac{S(T(t)) - \Phi(S(T(t))) - (I - P')D^{-1}T(t)}{\sqrt{r}}\right) + CY^r(t)$$

$$= W^r(0) + \Xi^r(t) + U^r(t)$$

where $W^r(0) = rM^r z$ for some $z \geq 0$,

$$\Xi^r(t) = rM^r \left(\frac{E(t) - \lambda t}{\sqrt{r}} - \frac{S(T(t)) - \Phi(S(T(t))) - (I - P')D^{-1}T(t)}{\sqrt{r}} \right) + \sqrt{r}(\rho^r - e)t,$$

and the second equality follows from $CY^r = \sqrt{r}CR^{r^{-1}}\lambda^r t - \sqrt{r}CT^r(t) = \sqrt{r}[et - CT^r(t)] + \sqrt{r}[CD^r(I - P')^{-1}\lambda^r - e]t$. The expression for Ξ^r explains the assumption that $\sqrt{r}(\rho^r - e)t \to \nu$.

Proposition 7.4. (Diffusion limit) Assume that $W^{r,*}(0) \Rightarrow w_0$, where w_0 is some properly defined random variable. Then,

$$(Z^{r,*}, W^{r,*}, U^{r,*}, \Xi^{r,*}) \Rightarrow (Z^*, W^*, U^*, \xi),$$

where (Z^*, W^*, U^*, ξ) satisfies (2.6)-(3.3) and $Z^* = \Delta(W^*)$, where $\Delta(\cdot)$ is defined in (3.5).

The proof of the large deviations estimates of Proposition 7.1 is given in Appendix B.

Proof of Proposition 7.2: Refreshing some notation, we have that $q^r(t) = Q^r(r^{-1/2}t)/\sqrt{r}$, $w^r(s) = Mq^r(s)$ $q^{r,\theta}(t) = [q^r(t) - \theta^r/\sqrt{r}]^+$ and $w^{r,\theta}(t) = [w^r(t) - (e - \rho^r)l^r\sqrt{r} - \eta^r/\sqrt{r}]^+$. We will use the shorthand notation Q^r for $Q^r(r^{-1/2}t)$ and q^r for $q^r(s)$. Expressions with argument Q^r are evaluated using the unscaled system dynamics with arrival and service rates given by λ^r and μ^r respectively, whereas expressions that use q^r use the fluid scaled dynamics with scaled arrival and service rate vectors given by $\lambda^r/\sqrt{r}$ and $\mu^r/\sqrt{r}$. For example, using (4.3), (4.4) and (4.9) we get that

$$\Psi^{\min}(l^r; Q^r, \Delta(W^r)) = r^{1/2}\Psi^{\min}(l^r\sqrt{r}; q^r, \Delta(w^r)), \tag{7.4}$$

$$T^{\min}(l^r; Q^r, \Delta(W^r)) = r^{-1/2}T^{\min}(l^r\sqrt{r}; q^r, \Delta(w^r)), \tag{7.5}$$

$$z(q^r) = r^{-1/2}z(Q^r). \tag{7.6}$$

Fluid scaled cumulative allocation process is given by $\bar{T}^{r,*}(\cdot) = r^{1/2}T^{r,*}(r^{-1/2}\cdot)$. This is uniformly Lipschitz and thus also a relatively compact family. Hence, there exists a converging subsequence $\{r_j\}$ with $r_j \to \infty$ such that $\bar{T}^{r_j,*} \Rightarrow \bar{T}^*$, where $\bar{T}^*$ is some limit process. Using the FSLLN and the key renewal theorem as in Theorem 4.1 in Dai [14] we conclude that $q^{r_j,*}$ will also converge to some limit trajectory q^*. Without loss of generality, we work directly with the converging subsequence, thus avoiding double subscripts. Then,

$$\begin{aligned}
\bar{T}^r(t) \;=\; \int_0^t dT^r(r^{-1/2}s) &\overset{(a)}{\approx} \int_0^t v(Q^r(r^{-1/2}s))ds \\
&= \int_0^t \frac{T^{\min}\left(l^r; Q^r(r^{-1/2}s), z(Q^r(r^{-1/2}s))\right)}{l^r}ds \\
&\overset{(b)}{=} \int_0^t \frac{T^{\min}\left(l^r\sqrt{r}; q^r(s), z(q^r(s))\right)}{l^r\sqrt{r}}ds \\
&\overset{(c)}{\approx} \int_0^t \frac{T^{\min}\left(l^r\sqrt{r}; q^{r,\theta}(s), \Delta(w^{r,\theta}(s))\right)}{l^r\sqrt{r}}ds
\end{aligned}$$

$$\stackrel{(d)}{\longrightarrow} \int_0^t \dot{T}^{\min}(0; q(s), \Delta(w(s)))\, ds$$
$$\stackrel{(e)}{=} \int_0^t dT^{\min}(s; z, \Delta(w)) \;=\; T^{\min}(t; z, w).$$

The remainder of the proof will justify steps (a)-(e). From Proposition 7.1 it follows that

$$\sum_r \mathbb{P}\left(\inf_{\kappa^r\theta^r/r \le s \le t} Q^r(s) \not\geq e \right) < \infty,$$

which implies from the Borel-Cantelli Lemma that for almost all sample paths $Q^r(s) > 0$ for all times $\kappa^r\theta^r/r \le s \le t$, as $r \to \infty$. Hence,

$$\left| \int_0^t dT^r(r^{-1/2}s) - \int_0^t v(Q^r(r^{-1/2}s))ds \right| \approx \kappa^r\theta^r/r \to 0.$$

This proves step (a). Step (b) follows from (7.5). Step (c) is established as follows. Note that

$$\frac{\left| q^r(\cdot) - q^{r,\theta}(\cdot) \right|}{l^r\sqrt{r}} \le \frac{\theta^r}{l^r r} = \frac{1}{c_l\theta^r} \;\to\; 0, \quad \text{as } r \to \infty.$$

Similarly, one gets that

$$\frac{\left| z(q^r(s)) - \Psi^{\min}(l^r\sqrt{r}; q^{r,\theta}(s), \Delta(w^{r,\theta}(s)) \right|}{l^r\sqrt{r}} \;\to\; 0, \quad \text{as } r \to \infty.$$

Hence, $\left| T^{\min}(l^r\sqrt{r}; q^r(s), z(q^r(s))) - T^{\min}(l^r\sqrt{r}; q^{r,\theta}(s), \Delta(w^{r,\theta}(s))) \right| / l^r\sqrt{r} \to 0$, as $r \to \infty$. Existence of the limit in (d) follows from [14, Theorem 4.1]. From (4.5), we have that $l^r\sqrt{r} \to 0$ as $r \to \infty$. It follows that $q^{r,\theta}(\cdot) \to q(\cdot)$ and $w^{r,\theta}(\cdot) \to w(\cdot) = Mq(\cdot)$. From (4.3) and (4.4) we see that for any fixed r, $T^{\min}(\cdot; z, z^\infty)$ is continuous in z and z^∞. Passing to the limit and using the continuous mapping theorem we get (d). Finally, (e) follows from the "straight line" nature of the minimum time trajectory and conditions (4.3) and (4.4). Steps (a)-(e) are all true almost surely. By the Lipschitz continuity of the allocation processes it follows that all of the above conditions are true for any s such that $0 \le s \le t$, which implies that the convergence is uniform on compact sets for all $t \ge 0$. This completes the proof. $\diamond$

Sketch of Proof of Proposition 7.3: This is a variation of Bramson's state space collapse result [7, Theorem 4]. Rather than reproducing a lengthy analysis, we will first point out the differences between his setup and ours, and then provide necessary arguments that would extend his work [7, §§4-8] to our case; it is useful to have a copy of [7] handy while going through this proof.

We start by making the following general remarks: (i) all probabilistic assumptions imposed in [7, §§2-3] are also in place here; (ii) following the footnote of §5.1, the scalings used here and those used in [7] are equivalent; (iii) the head-of-line property used in [7] also holds here; (iv) the definition of workload used in [7] is somewhat different from the one used here, but our version only simplifies matters by making W just a linear combination of the queue length vector, $W = MQ$; (v) the main differences that we will need to account for are: (a) our policy may idle a server when idleness could be prevented, whereas Bramson restricts attention to non-idling rules and (b) the function Δ that maps workload to queue lengths is not linear as in Bramson, but it is well defined and continuous. The non-idling restriction that appears to be the biggest difference between the two setups is not crucial. Indeed, Bramson's restriction to non-idling policies was a logical starting step but it is not essential in proving the desired result.

First, note that the results of [7, §4] are still valid since all probabilistic assumptions of [7] are also in place here. The results of [7, §5] on WLLN for the arrival, service and routing processes and for the almost Lipschitz continuity of fluid scaled network processes also go through unchanged. These results only assume that the policy is HL and non-idling. Careful inspection of the proof, however, reveals that the latter is not used. It remains to establish the equivalents of the results presented in [7, §6] for tracking policies. The reader is referred to [7, §7] for an overview of these results.

Proposition 6.1 of [7] says that fluid scaled network processes are close to Lipschitz cluster points for large enough r. This is a consequence of [7, Prop. 4.1] that does not require any specific policy assumptions, and goes through unchanged in our setup. Proposition 6.2 of [7] concludes that the cluster points specified above are solutions of the fluid model equations given in our Proposition 7.2. Bramson's proof goes through by replacing his fluid model equations by the ones derived in Proposition 7.2 that are satisfied by the fluid limits under our tracking policy.

Proposition 6.3 in [7] is policy specific. It establishes the property of "uniform convergence" for the associated fluid model; see the discussion in Bramson and Dai [8, §5]. Uniform convergence requires that starting from an arbitrary initial condition $q(0) = z$ with $|z| \leq 1$, and setting the target state to $q(\infty) = \Delta(w_0)$ where $w_0 = Mz$ is the initial workload position, there exists a fixed function $H(t)$ such that $|q(t) - q(\infty)| \leq H(t)$, and $H(t) \to 0$ as $t \to \infty$. In the proof of [7, Theorem 4] this is replaced by [7, Assumption 3.1]. Here, we prove directly the equivalent to [7, Proposition 6.3], by constructing the appropriate function $H(t)$ that will establish uniform convergence. This step is easy. By observing the fluid model equations derived in Proposition 7.2, we get that starting from any $|z| \leq 1$, the fluid model is guaranteed to have reached its target $q(\infty)$ by time t^*, where

$$t^* = \max\left\{t^{\min}(z, \Delta(w_0)) : |z| \leq 1,\ w_0 = Mz\right\}.$$

Setting $H(t) = t^{\min}(q(t), \Delta(w(t)))$ will do.

Proposition 6.4 in [7] follows after three changes. First, the function $H(t)$ used by Bramson should be replaced by the one identified above. Second, [7, eq.(6.40)] that depends on the definition of workload needs to be changed. In our case, $W(t) = Mq(t)$, and the condition equivalent to [7, eq.(6.40)] follows directly from Proposition 6.3. Finally, [7, eq.(6.41)] uses the norm of the mapping Δ, which is denoted by B_{15}. In our setting, we have that for $z^* = \Delta^r(w)$, $|z^*| \leq |w| \|\Delta^r\|$, where $\|\Delta^r\| = \min_k rm_k^r$, $\min_k rm_k^r \to \min_k m_k$ as $r \to \infty$, and $m_k = 1/\mu_k$ is the limiting mean service time for each class k. Plugging in for B_{15} and using the continuity of Δ we complete the proof of Proposition 6.4. (Note in passing that the linearity of Δ is not needed, and continuity is sufficient.)

Corollary 6.2 of [7] follows from [7, Propositions 5.2, 6.1-6.4], and Lemmas 6.2-6.3 and Proposition 6.5 of [7] do not require any specific policy assumptions. All these results still apply in our setting. The proof of Theorem 4 in [7] follows as a consequence of [7, Proposition 6.5] and [7, Corollary 5.1]. This also completes the proof of our proposition. $\diamond$

Lemma 1. $T^{r,*}(t) \Rightarrow R^{-1}\lambda t = \alpha t$ for $t \geq 0$.

Proof: $\{T^{r,*}\}$ is uniformly Lipschitz and thus also a relatively compact family. Hence, there exists a subsequence $\{r_n\}$ with $r_n \to \infty$ such that $T^{r_n,*} \Rightarrow T^*$, where T^* is some limit process. We consider a "double fluid scaling" and analyze the limits of

$$\tilde{q}^{r_n} = \frac{Q^{r_n}}{r_n} \quad \text{and} \quad \tilde{u}^{r_n} = \frac{I^{r_n}}{r_n}.$$

From the renewal theorem we have that

$$\frac{1}{r_n}E^{r_n}(t) \Rightarrow \lambda t, \quad \frac{1}{r_n}S^{r_n}(T^{r_n,*}(t)) \Rightarrow D^{-1}T^*(t), \quad \text{and} \quad \frac{1}{r_n}\Phi(S^{r_n}(T^{r_n,*}(t))) \Rightarrow P'D^{-1}T^*(t).$$

From Proposition 7.2 we have that

$$\tilde{q}(t) = 0 + \lambda t - RT^*(t), \quad T^*(0) = 0 \quad \text{and} \quad \tilde{q}(\cdot) = \Psi^{\min}(\cdot; 0, \Delta(w_0)).$$

For $\tilde{q}(0) = 0$, we have that $w_0 = 0$, $\Delta(w_0) = 0$, and thus $\tilde{q}(t) = 0$ for all $t \geq 0$. It follows that $T^*(t) = R^{-1}\lambda t$ for all $t \geq 0$. $\diamond$

Proof of Proposition 7.4: This is a direct application of Williams' invariance principle [46, Theorem 4.1]. To apply that result we need to rewrite the system equations in the form

$$W^{r,*}(t) = W^r(0) + \Xi^{r,*}(t) + U^{r,*}(t), \tag{7.7}$$

$$U^{r,*}(t) = \tilde{U}^{r,*}(t) + \gamma^{r,*}(t) \text{ and for some constant } \delta^{r,*} \geq 0, \text{ for all servers } i, \tag{7.8}$$

$$(a)\ \tilde{U}^{r,*}(0) = 0 \quad \text{and} \quad \tilde{U}^{r,*} \text{ is non-decreasing,} \tag{7.9}$$

$$(b)\ \int_0^\infty \mathbf{1}_{\{W_i^{r,*}(s) > \delta^{r,*}\}} d\tilde{U}_i^{r,*}(s) = 0, \tag{7.10}$$

and $\gamma^{r,*}, \delta^{r,*} \to 0$ in probability as $r \to \infty$. The interpretation of these conditions is that the triple (W^r, U^r, Ξ^r) "almost" satisfies the limiting equations of the optimally controlled Brownian model, and asymptotically all error terms become negligible.

Using Lemma 1, the probabilistic assumptions for the arrival and service time processes and using a random time change argument we get that $\Xi^{r,*}(t) \Rightarrow \xi(t)$, where ξ is the S-dimensional Brownian motion of §2.2; see for example [45, Theorem 7.1]. Decompose the scaled idleness process in the form

$$U_i^{r,*}(t) = \int_0^t \mathbf{1}_{\{W_i^{r,*}(s) > \delta^{r,*}\}} dU_i^{r,*}(s) + \int_0^t \mathbf{1}_{\{W_i^{r,*}(s) \leq \delta^{r,*}\}} dU_i^{r,*}(s),$$

and set $\delta^{r,*} = \eta^r/\sqrt{r} \to 0$, as $r \to \infty$. The first term accounts for all idleness incurred while the workload process $W_i^r > \delta^{r,*}$ and the second term accounts for idleness incurred when $W_i^r \leq \delta^{r,*}$. In the above expression, we set the first term equal to $\gamma^{r,*}(t)$ and the second to $\tilde{U}^{r,*}(t)$. Hence, (7.10) is satisfied. It remains to show that in the heavy traffic limit, $\gamma^{r,*}(t) \to 0$ in probability.

If at time t, $w_i = (Mq)_i \geq \eta_i$, then we argue that $(Cv(q))_i = 1$; that is, the server will not idle intentionally. Recall that $w_i^\dagger = w_i - (1-\rho_i)l$, $(Mz^*)_i = w_i^\dagger$, and $(Mz(q))_i = w_i^\dagger$. By construction of the minimum time trajectory $\Psi^{\min}$, we have that $\Psi^{\min}(l; q, z(q)) = z(q)$. Hence, at time l along this trajectory, the total workload for server i will be given by $(Mz(q))_i = w_i^\dagger$. Using (4.11) we see that with minimum time control the server cannot incur any idleness up to time l, or equivalently, that $(Cv(q))_i = 1$. In this case, the server i can only idle if for some $k \in C_i$, $v_k(q) > 0$, but $q_k = 0$. From Proposition 7.1, $\mathbb{P}(Q^{r,*}(t) \not\geq e) \leq c/r^2 \to 0$, as $r \to \infty$. Hence,

$$\mathbb{E}\gamma^{r,*}(t) \leq \kappa^r l^r + t\frac{t}{r^2} \to 0,$$

which implies that $\gamma^{r,*} \to 0$, in probability.

Applying Williams' invariance principle [46, Theorem 4.1, Proposition 4.2], using the continuity of Δ, the state space collapse condition established in Proposition 7.3, and the fact that the reflection matrix, which is just the identity matrix, is *completely-$\mathcal{S}$* and of the form $I + Q$ for some matrix Q, we complete the proof. $\diamond$

Proof of Theorem 1: The proof is divided in two parts. First, the LHS inequality is established by considering any converging subsequence that achieves the "lim" in the "liminf" and invoke the pathwise optimality of the optimal Brownian control. We follow Bell and Williams [4, §9, Thm. 5.3].

Consider any sequence of policies $T = \{T^r\}$, one for each set of parameter values, and denote by $J(T^r)$ the corresponding cost achieved under this policy. The first part of the proof will establish that $\liminf_{r\to\infty} J^r(T^r) \triangleq \underline{J}(T) \geq J^*$. We only consider sequences $\{T^r\}$ for which $\underline{J}(T) < \infty$, otherwise our claim trivially follows. It is easy to characterize the set of such sequences that need to be considered. Imitating the approach of [4, Lemma 9.3]. we get that it must be that as $r \to \infty$, $T^r(t) \Rightarrow \alpha t$, for all $t \geq 0$. We argue by contradiction. Consider any subsequence $\{T^{r'}\}$ along which the "liminf" in $\underline{J}(T)$ is achieved. By the FCLT for the arrival and service time processes we have that $r'^{-1/2}\chi^{r'}(t) \Rightarrow X(t)$, where $X(t)$ is the K-dimensional Brownian motion described in §2. Suppose that $T^{r'}(t) \not\Rightarrow \alpha t$. Specifically, assume that at time $t = t^*$, $\lim_{r'} T^{r'}(t^*) \not\Rightarrow \alpha t^*$ This implies that $\lim_{r'} r'^{-1/2} R^{r'} \delta^{r'}(t^*) = \lim_{r'} r'^{-1/2} R^{r'}(\alpha^r t^* - T^{r'}(t^*)) \to \infty$, and therefore that

$$Z^{r'}(t^*) = z^{r'} + r'^{-1/2}\chi^{r'}(t^*) + r'^{-1/2} R^{r'} \delta^{r'}(t^*) \to \infty, \quad \text{as } r' \to \infty,$$

which contradicts the finiteness of $\underline{J}(T)$. Hence, it must be that $T^{r'}(t) \Rightarrow \alpha t$, for all $t \geq 0$.

We proceed by observing that

$$\begin{aligned} W^r(t) &= W^r(0) + \Xi^r(t) + U^r(t) \\ &\geq W^r(0) + \Xi^r(t) + \left[-\inf_{0\leq s\leq t} W^r(0) + \Xi^r(s) \right]^+ \triangleq \phi(W^r(0) + \Xi^r(t)). \end{aligned}$$

This minimality and continuity of the mapping ϕ are well known (see Harrison [18, §2.2]).

Given that $T^r(t) \Rightarrow \alpha t$, and the fact that this limit is deterministic we get that

$$(T^{r'}(\cdot), \Xi^{r'}(\cdot)) \Rightarrow (\alpha\cdot, \xi(\cdot)).$$

Invoking the Skorokhod representation theorem, we may choose an equivalent distributional representation (using the same symbols) such that all processes together with their limits are defined on the same probability space and weak convergence is replaced by almost sure u.o.c. convergence of the corresponding sample paths. In this case, for any sequence T such that $\underline{J}(T) < \infty$ consider any subsequence $\{T^{r'}\}$ along which the "liminf" is achieved. By Fatou's Lemma we have that

$$\underline{J}(T) = \lim_{r' \to \infty} J^{r'}(T^{r'}) \geq \mathbb{E}\left(\int_0^t e^{-\gamma t} \liminf_{r' \to \infty} \sum_k c_k Z_k^{r'}(t) dt \right).$$

Fix any sample path ω for which the above a.s. convergence holds and pick any subsequence $\{T^{r''}\}$ that achieves the "liminf" in the RHS of the above equation. Then,

$$\lim_{r'' \to \infty} W^{r''}(t,\omega) \geq \phi(w_0 + \xi(t,\omega)) = W^*(t,\omega),$$

which implies that $\lim_{r'' \to \infty} Z^{r''}(t,\omega) \geq \Delta(W^*(t)) = Z^*(t,\omega)$ and $\underline{J}(T) \geq J^*$.

Second, we want to establish the RHS of (5.5). This is done in two steps. First, we show that $\{Z^{r,*}\}$ is uniformly integrable (UI) (as in [29, §7] or [4, §9]), and then pass to the limit and take expectations for the limit processes. Given that $W^r = rM^r Z^r$, it suffices to show that for any fixed t, the family $\{\sup_{0 \leq s \leq t} W^{r,*}(s)\}$ is UI.

Given (7.7)-(7.10) we can apply Williams' oscillation inequality [46, Theorem 5.1], that bounds the variation of the workload process in terms of the variation of Ξ^r, γ^r and δ^r, to get that

$$\sup_{0 \leq s \leq t} |W^{r,*}(s)| \;\leq\; |W^r(0)| + (S+1) \sup_{0 \leq s \leq t} |\Xi^{r,*}(s)| + (S+1)\left(\gamma^{r,*}(t) + \delta^{r,*}\right); \tag{7.11}$$

this is identical to [45, eq.(39)]. The constant $(S+1)$ that appears in this expression, comes out of Williams's work. The appropriate value depends only on the reflection matrix. In our case, this is the identity matrix, and the constant is equal to the number of servers in the system plus one.

We need to establish that $\{\Xi^{r,*}\}$ and $\{\gamma^{r,*}\}$ are UI. Under the probabilistic assumptions regarding the arrival and service time processes it is easy to show that $\Xi^{r,*}(t)$ is UI; e.g., following the proofs in [29, Theorem 7] or [4, Theorem 5.3]. Mimicking the derivation in Proposition 7.4, we get that

$$\mathbb{E}(\gamma^{r,*}(t))^2 \leq rt^2 \frac{t}{r^2}$$

which in turn implies that $\{\gamma^{r,*}\}$ is also UI. (The structure of our policy is used in proving that $\gamma^{r,*}$ is UI, which corresponds to establishing the UI property for the idleness process.)

Hence, $\{\sup_{0 \leq s \leq t} Z^{r,*}(s)\}$ is a UI family. Passing to the limit we get that

$$\begin{aligned}
\lim_{r \to \infty} \mathbb{E}\left(\int_0^\infty e^{-\gamma t} \sum_k c_k Z^{r,*}(t) dt \right) &= \mathbb{E}\left(\int_0^\infty e^{-\gamma t} \lim_{r \to \infty} \sum_k c_k Z^{r,*}(t) dt \right) \\
&= \mathbb{E}\left(\int_0^\infty e^{-\gamma t} \sum_k c_k Z^*(t) dt \right) = J^*,
\end{aligned}$$

where the first equality follows from UI property, the second from the diffusion limit derived in Proposition 7.4, and the optimality from the properties of Z^* given in §3. ⋄

Proof of Theorem 2: The proof is identical to that of Theorem 1, subject to two changes. The first is in establishing condition (d) in Proposition 7.2. This will follow from the arguments of [34, Theorem 5.1]. The second is in identifying the function $H(t)$ needed in Proposition 7.3. For the problem in (6.2), one can easily show that the terminal value of minimum cost will be achieved in finite time. The analysis of [34, Proposition 6.1] can be modified to produce a bound on that time-to-target that can be used in place of $H(t)$. $\diamond$

Appendix B: Large deviations bounds of Proposition 7.1

We provide an outline of the proof of this Proposition as most of the details can be found in the book of Shwartz and Weiss [42, §6.1]. We start with an intuitive sketch of the argument, and then provide a more detailed roadmap, giving appropriate references from [42, §6.1].

To adhere to the notation used in [42], rather than using our scaling where $\lambda^r \to \infty$ and time stays unscaled, we will stretch out all random variables such that the new arrival and service rates will be $\hat{\lambda}^r = \lambda^r/r$ and $\hat{\mu}^r = \mu^r/r$ but time will be stretched out by a factor r. The two setups are identical. Specifically, denoting the queue length process for that augmented system by $\hat{Q}^r(\cdot)$, and setting $\hat{\theta}^r = c_\theta \log(r)$ and the planning horizon $\hat{l}^r = c_l {\theta^r}^2$, we have that for any $\epsilon > 0$,

$$\mathbb{P}\left(\inf_{\epsilon\le s\le t} Q^{r,*}(s) \not\geq e\right) \;=\; \mathbb{P}\left(\inf_{r\epsilon\le s\le rt} \hat{Q}^{r,*}(s) \not\geq e\right), \tag{7.12}$$

and hence it is sufficient to get a bound for the RHS probability.

Elementary large deviations results, like the ones in (5.2)-(5.3), bound the probability of observing large fluctuations away from expected behavior of a sample (or time) average, where the size of this fluctuation is proportional to the sample size (or observation period). On the other hand, the bound in (7.12) is concerned with the probability of seeing fluctuations of order $\hat{\theta}^r$ over long observation periods of length $r \gg \hat{\theta}^r$. Its proof will follow a standard procedure for results of that short. First, one bounds the probability that a queue length gets depleted over a period of length $\hat{\theta}^r$, for which a simple large deviations argument will suffice. Then, many such intervals, $rt/\hat{\theta}^r \triangleq N$ to be precise, are pieced together to make up the longer period of length r. The relevant question now is how many such intervals will it take until one of the queue lengths does get depleted. This step follows a standard Friedlin-Wentzell calculation of an exit time from some domain of interest, which here is the region $\{q : q \geq e\}$; this part will follow [42, §6.1]. Finally, we need to take care of the initial condition, and specifically, allowing for the system to start empty, we need to bound the time it takes to reach close to $\hat{\theta^r}$. Using the first two steps one can then infer that once the system has built up enough safety stock for each queue, this will never get depleted again; this will take care of problems that arise due to boundary behavior.

The first and third steps are summarized in the following lemmas; their proofs are given at the end of the appendix.

Lemma 2. For r sufficiently large,

$$\mathbb{P}\left(\inf_{0\leq s\leq \theta^r/r} Q^r(s) \not\geq e \mid Q^r(0) \geq \theta^r/2\right) \leq ae^{-c_{.5}\theta^r}, \tag{7.13}$$

for some constants $a, c_{.5} > 0$.

Lemma 3. Let $\tau^{0\to .5\theta^r} = \inf\{t : Q^r(t) \geq \theta^r/2;\ Q^r(0) < \theta^r/2\}$. For r sufficiently large and for some appropriate (finite) constant κ^r,

$$\mathbb{P}\left(\tau^{0\to .5\theta^r} > \kappa^r l^r\right) \leq \frac{t}{2r^2}. \tag{7.14}$$

The second part of the proof follows [42, §6.1] and will not be reproduced here. We will only trace the steps of the proof and provide appropriate pointers to the results in [42, §6.1]. First, we note that within each of the N sub-intervals system behavior can be analyzed using Lemma 2. Let $\tau^{.5\theta^r\to 0} = \inf\{t : Q^r(t) \not\geq e;\ Q^r(0) \geq \theta^r/2\}$ be the first time until some queue length gets depleted. From (7.13) one would expect that within each sub-interval (a) it is very unlikely that some queue length does get depleted, and (b) it is very likely that the queue length process will be above $\delta\theta^r$, for some $0 < \delta < 1/2$. Indeed the derivation of Lemma 2 will also give us that

$$\mathbb{P}\left(Q^r(\theta^r/r) \geq \delta\theta^r \mid Q^r(0) \geq \delta\theta^r\right) \geq 1 - ae^{-c'_\delta\theta^r}. \tag{7.15}$$

for some constant $c'_\delta > 0$; this is [42, Lemma 6.32]. Similarly, a simple modification of Lemma 2 yields the following bound

$$\mathbb{P}\left(\inf_{i\theta^r/r\leq s\leq (i+1)\theta^r/r} Q^r(s) \not\geq e \mid Q^r(i\theta^r/r) \geq \delta\theta^r\right) \leq ae^{-c_\delta\theta^r}. \tag{7.16}$$

where $0 < c_\delta < c_{.5}$; this is the result in [42, Ex. 6.34]. Thus, consecutive sub-intervals would amount to roughly independent attempts for some of the queue lengths to empty. Intuitively speaking, the queue length process will not drain slowly until it empties some buffer, but it will do so in a sudden excursion away from its mean field behavior. Therefore, one would expect that $\mathbb{E}\tau^{.5\theta^r\to 0} \approx e^{c_\delta\theta^r}/a$, or equivalently, that it will take $e^{c_\delta\theta^r}/a\theta^r$ sub-intervals of length θ^r until the first exit time from our domain of interest, $\tau^{\theta^r\to 0}$. The bound of interest is obtained by calculating the probability of a successful excursion outside $\{q : q \geq e\}$ in N independent attempts that is given by

$$\mathbb{P}\left(\tau^{.5\theta^r\to 0} < rt\right) \approx Nae^{-c_\delta\theta^r} \leq ate^{-c_\delta\theta^r+\log(r)}. \tag{7.17}$$

Set $\epsilon = \kappa^r l^r$ in (7.12) and fix $\delta \in (0, 1/2)$. Choosing the constant c_θ in (4.5) such that $c_\theta > 3/c_\delta$, would give us (7.12). Combining with (7.14) we would get (7.2) and complete the proof of the Proposition. The actual proof of this result follows the first part of the proof of Theorem 6.17 in [42] and uses the results of exercises 6.41 and 6.42 of [42] and the bounds in (7.15) and (7.16). To avoid duplication of a long treatment, the reader is referred to their text for the detailed argument.

Proof of Lemma 2. We consider directly the rescaled process $\hat{Q}^r(\cdot)$, and assuming that $\hat{Q}^r(0) \geq \theta^r/2$, we want to analyze the system behavior over period of length θ^r; recall that this corresponds to θ^r/r time units for the Q^r process. In the sequel we treat r as fixed.

Our approach studies an "augmented" network with arrival rates $\tilde{\lambda} = \hat{\lambda}^r$ and service rates $\tilde{\mu} = \hat{\mu}^r$, and considers the tracking policy defined in §4 for $\tilde{\theta} = |\lambda|$ and $\tilde{l} = c_l\tilde{\theta}^2$. The queue length and cumulative allocation processes of such a system will be denoted by $\tilde{X}(\cdot)$ and $\tilde{T}(\cdot)$. We will then consider fluid scaled processes according to $\bar{\tilde{X}}^n(t) = \tilde{X}^n(nt)/n$ and $\bar{\tilde{T}}^n(t) = \tilde{T}(nt)/n$, and an initial condition $\bar{\tilde{X}}^n(0) = \frac{1}{2}e$. It is easy to see that $(\hat{Q}^r, \hat{T}^r)$ is equal to $(\bar{\tilde{X}}^n, \bar{\tilde{T}}^n)$ for $n = \theta^r$.

It is sufficient to show that for some positive constants $a, c_{.5} > 0$,

$$\mathbb{P}\left(\inf_{0\le s\le 1} \bar{\tilde{X}}^n(s) \not\ge e\right) \le ae^{-c_{.5}n}.$$

The result in (7.13) will then follow by setting $n = \theta^r$.

Abusing earlier notation, we will denote the cumulative arrival and service time processes by $\tilde{E}$ and $\tilde{V}$, and let $\bar{\tilde{E}}^n(t) = \tilde{E}(nt)/n$ and $\bar{\tilde{V}}^n(t) = \tilde{V}(nt)/n$ -these processes have rates $\tilde{\lambda}$ and $\tilde{\mu}$ respectively. Finally, $\bar{\tilde{S}}^n(\tilde{T}(t)) = \tilde{S}(\tilde{T}(nt))/n$.

The equations of dynamics are given by

$$\bar{\tilde{X}}^n(t) = \bar{\tilde{X}}^n(0) + \bar{\tilde{E}}^n(t) - \bar{\tilde{S}}^n(\tilde{T}(t)) + \sum_{k=1}^{K} \tilde{\Phi}^{k,n}(S^n(\tilde{T}(t))). \tag{7.18}$$

From (5.2)-(5.3) we have that

$$\mathbb{P}\left(\sup_{0\le s\le 1} |\bar{\tilde{E}}^n(s) - \tilde{\lambda}s| > \frac{\epsilon}{6}\right) \le c_a e^{-f_a(\epsilon/6)n}, \tag{7.19}$$

where $f_a^r(\epsilon/6) = \min_k f_a^{r,k}(\epsilon/6K)$ is the exponent when the arrival rates are λ^r in the original system. For $\bar{\tilde{S}}^n(\tilde{T}(t))$ we have that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\le s\le 1} |\bar{\tilde{S}}^n_k(\tilde{T}(s)) - \tilde{\mu}_k\bar{\tilde{T}}^n_k(s)| > \frac{\epsilon}{6K}\right) = \\
&= \mathbb{P}\left(\sup_{0\le s\le 1} \bar{\tilde{S}}^n_k(\tilde{T}(s)) < \tilde{\mu}_k\bar{\tilde{T}}^n_k(s) - \frac{\epsilon}{6K}\right) + \mathbb{P}\left(\sup_{0\le s\le 1} \bar{\tilde{S}}^n_k(\tilde{T}(s)) > \tilde{\mu}_k\bar{\tilde{T}}^n_k(s) + \frac{\epsilon}{6K}\right) \\
&= \mathbb{P}\left(\sup_{0\le s\le 1} \frac{1}{n}\tilde{V}^n_k\left(\tilde{\mu}_k\tilde{T}_k(ns) - \frac{n\epsilon}{6K}\right) > \bar{\tilde{T}}^n_k(t)\right) + \mathbb{P}\left(\sup_{0\le s\le 1} \frac{1}{n}\tilde{V}^n_k\left(\tilde{\mu}_k\tilde{T}_k(ns) + \frac{n\epsilon}{6K}\right) < \bar{\tilde{T}}^n_k(t)\right) \\
&= \mathbb{P}\left(\sup_{0\le s\le 1} \frac{1}{n}\tilde{V}^n_k\left(\tilde{\mu}_k\tilde{T}_k(ns) - \frac{n\epsilon}{6K}\right) - \left(\bar{\tilde{T}}^n_k(t) - \frac{\hat{m}^r_k\epsilon}{6K}\right) > \frac{\hat{m}^r_k\epsilon}{3K}\right) \\
&\qquad + \mathbb{P}\left(\sup_{0\le s\le 1} \frac{1}{n}\tilde{V}^n_k\left(\tilde{\mu}_k\tilde{T}_k(ns) + \frac{n\epsilon}{6K}\right) - \left(\bar{\tilde{T}}^n_k(t) + \frac{\hat{m}^r_k\epsilon}{6K}\right) < -\frac{\hat{m}^r_k\epsilon}{6K}\right) \\
&\le 2e^{-f_s^{k,r}(\hat{m}^r_k\epsilon/6KT_s)n},
\end{aligned}$$

where the last bound follows from (5.2)-(5.3) and $T_s = \max_k \tilde{\mu}_k + 1/6K$. Summing over all k, for $f_s^r(\epsilon/3) = \min_k f_s^{r,k}(\hat{m}^r_k\epsilon/6KT_s)$ and some constant $c_s > 0$,

$$\mathbb{P}\left(\sup_{0\le s\le 1} \left|\bar{\tilde{S}}^n(\tilde{T}(s)) - D^{-1}\bar{\tilde{T}}^n(s)\right| > \frac{\epsilon}{6}\right) \le c_s e^{-f_s(\epsilon/6)n}, \tag{7.20}$$

A similar argument will establish that

$$\mathbb{P}\left(\sup_{0\le s\le 1}\left|\sum_{k=1}^{K}\bar{\bar{\Phi}}^{k,n}(S^n(\tilde{T}(s))) - P'D^{-1}\bar{T}^r(s)\right| > \frac{\epsilon}{6}\right) \le c_r e^{-f_r(\epsilon/6)r}, \tag{7.21}$$

where $c_r > 0$ is some appropriate constant and $f_r(\epsilon/6)$ is again obtained as the minimum over all exponents $f_r^{kl}(a_{kl}\epsilon/6KT_r)$, where a_{kl} and T_r depend on the vector m and the routing matrix P.

Summing (7.19)-(7.21) and for some constant $c > 0$ and r sufficiently large ($r \ge r^*$),

$$\mathbb{P}\left(\sup_{0\le s\le 1}\left|\bar{\bar{X}}^n(s) - \bar{\bar{X}}^n(0) + \tilde{\lambda}s - R\bar{\bar{T}}^n(s)\right| > \epsilon/2\right) \le a e^{-\min\{f_a(\epsilon/6), f_s(\epsilon/6), f_r(\epsilon/6)\}n}. \tag{7.22}$$

It remains to analyze $\bar{\bar{X}}^n(0) + \tilde{\lambda}s - R\bar{\bar{T}}^n(s)$. We do this by comparing it to its fluid limit obtained as $n \to \infty$. By Theorem 4.1 of Dai [14] we know that $(\bar{\bar{X}}^n(t), \bar{\bar{T}}^n(t)) \to (\bar{\tilde{X}}(t), \bar{\tilde{T}}(t))$ u.o.c.. Therefore, for any $\epsilon' > 0$ there exists a large enough $N(\epsilon')$ such that for $n > N(\epsilon')$ and for all $t \ge 0$, $|\bar{\bar{T}}^n(t) - \bar{\tilde{T}}(t)| < \epsilon'$. Then we have that

$$\sup_{0\le s\le 1}\left|(\bar{\bar{X}}^n(0) + \tilde{\lambda}s - R\bar{\bar{T}}^n(s)) - (\bar{\tilde{X}}(0) + \tilde{\lambda}s - R\bar{\tilde{T}}(s))\right| < \max_k \tilde{\mu}_k \epsilon'. \tag{7.23}$$

It is easy to see that $\bar{\tilde{X}}(0) = \frac{1}{2}e$. Condition (4.10) implies that $q + (\lambda - Rv(q))t^{\min}(q, z(q)) \ge \theta$, and $t^{\min}(q, z(q)) \le 2\tilde{l}$; the last assertion follows from the definition of l and (4.9). Translating this inequality to the variables considered here we get that $\bar{\bar{X}}^n(t) + (\tilde{\lambda} - Rv(\bar{\bar{X}}^n(t)))2\tilde{l}^n \ge e$. Note that $\dot{\bar{\bar{T}}}^n_k(t) = v_k(\bar{\bar{X}}^n(t))$ for all k such that $\bar{\bar{X}}^n_k(t) > 0$, and $\dot{\bar{\tilde{T}}}_k(t) = 0$, otherwise. It follows that $\bar{\bar{X}}^n(t) + (\tilde{\lambda} - R\dot{\bar{\bar{T}}}^n(t))2\tilde{l}^n \ge e$. Passing to the limit and using standard fluid limiting arguments (see Dai [14, §4-5]) we get that $\bar{\tilde{X}}_k(t) \Rightarrow (\tilde{\lambda} - R\dot{\bar{\tilde{T}}}(t))_k \ge 0$, for all $t \ge 0$. Given that $\bar{\tilde{X}}(0) = \frac{1}{2}e$, this implies that $\bar{\tilde{X}}(t) \ge \frac{1}{2}e$ for all $t \ge 0$. Set $\epsilon' = \epsilon/(2\max_k \tilde{\mu}_k)$. From (7.23), $\bar{\bar{X}}^n(0) + \tilde{\lambda}t - R\bar{\bar{T}}^n(t) \ge (1/2 - \epsilon/2)e$. Pick any $\epsilon < 1/2$ and set $c_{.5} = f(\epsilon) = \min\{f_a(\epsilon/6), f_s(\epsilon/6), f_r(\epsilon/6)\}$. Now, (7.13) follows from (7.22). $\diamond$

Proof of Lemma 3. The proof follows along the lines of Lemma 2. Consider the scenario where class k starts empty (this is the worse case scenario for each individual class). At time $t = 0$, the tracking policy of §4 guarantees a nominal drift for class k jobs given by

$$\forall\ k:\ Q_k^r = 0 \ \Rightarrow\ (\lambda^r - Rv(Q^r))_k \ \ge\ \frac{\theta^r}{t^{\min}(Q^r, z(Q^r))}.$$

That is, under this instantaneous allocation the content of the class k buffer will be above θ^r within $t^{\min}(Q^r, z(Q^r))$ time units. By construction of the target state $z(Q^r)$ it follows that $t^{\min}(Q^r, z(Q^r)) \le l^r + t^{\min}(\mathbf{0}, \theta^r) \triangleq a^r l^r$, where the constant a^r depends on λ^r and μ^r and can be computed through (4.1). This drift decreases as Q_k^r approaches θ^r, however,

$$\forall\ k:\ Q_k^r \le \delta\theta^r \ \Rightarrow\ (\lambda^r - Rv(Q^r))_k \ \ge\ \frac{\theta^r}{2t^{\min}(Q^r, z(Q^r))} \ \ge\ \frac{\delta\theta^r}{a^r l^r}. \tag{7.24}$$

This yields the following bound:

$$\mathbb{E}\tau^{0\to .5\theta^r} \le a^r l^r. \tag{7.25}$$

Proceeding as in Lemma 2 and using (7.24) to lower bound the cumulative allocation process up to time $\tau^{0\to .5\theta^r}$, and the bound in (7.25) we get that

$$\mathbb{P}\left(\tau^{0\to .5\theta^r} > \kappa^r l^r\right) \le \frac{t}{r^2},$$

where $\kappa^r > a^r$ is an appropriate constant that depends on the large deviations exponents for the arrival and service time random variables with rates λ^r and μ^r respectively. (One can think of κ^r as the appropriate multiple of a^r that will give the correct constant in the RHS of the large deviations bound.) This is a finite constant for each r, and thus $\kappa^r l^r \to 0$ as $r \to \infty$. $\diamond$

References

[1] F. Avram, D. Bertsimas, and M. Ricard. Fluid models of sequencing problems in open queueing networks; an optimal control approach. In F. Kelly and R. Williams, editors, *Stochastic Networks*, volume 71 of *Proceedings of the IMA*, pages 199–234. Springer-Verlag, New York, 1995.

[2] S. Axsäter. Continuous review policies for multi-level inventory systems with stochastic demand. In S. C. Graves, A. H. G. R. Kan, and P. H. Zipkin, editors, *Logistics of Production and Inventory*, volume 4 of *Handbooks in OR & MS*, pages 175–197. North-Holland, 1993.

[3] N. Bäuerle. Asymptotic optimality of tracking policies in stochastic networks. 1999. Preprint, University of Ulm, Germany.

[4] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: asymptotic optimality of a threshold policy. *Ann. Appl. Prob.*, 1999. Submitted.

[5] V. Böhm. On the continuity of the optimal policy set for linear programs. *SIAM J. on Applied Mathematics*, 28(2):303–306, 1975.

[6] M. Bramson. Stability of two families of queueing networks and a discussion of fluid limits. *Queueing Systems*, 28:7–31, 1998.

[7] M. Bramson. State space collapse with applications to heavy-traffic limits for multiclass queueing networks. *Queueing Systems*, 30:89–148, 1998.

[8] M. Bramson and J. Dai. Heavy traffic limits for some queueing networks. 1999. Preprint.

[9] H. Chen. Fluid approximations and stability of multiclass queueing networks: work-conserving policies. *Ann. Appl. Prob.*, 5:637–655, 1995.

[10] H. Chen and D. Yao. Dynamic scheduling of a multiclass fluid network. *Oper. Res.*, 41(6):1104–1115, Dec. 1993.

[11] M. Chen, C. Pandit, and S. Meyn. In search of sensitivity in network optimization. 2002. Submitted for publication.

[12] D. Clarke. *Advances in Model Predictive Control*. Oxford University Press, 1994.

[13] J. Dai and G. Weiss. A fluid heuristic for minimizing makespan in job-shops. *Oper. Res.*, 2002. To appear.

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

[15] A. Dembo and T. Zajik. Large deviations: from empirical mean and measure to partial sum processes. *Stoc. Proc. and Appl.*, 57:191–224, 1995.

[16] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. 2nd ed., Applications of Mathematics, Springer Verlag, 1998.

[17] P. Dupuis and K. Ramanan. A multiclass feedback queueing network with a regular Skorokhod problem. 1999. Preprint.

[18] J. M. Harrison. *Brownian motion and stochastic flow systems*. John Wiley & Sons, 1985.

[19] J. M. Harrison. Brownian models of queueing networks with heterogeneous customer populations. In W. Fleming and P. L. Lions, editors, *Stochastic Differential Systems, Stochastic Control Theory and Applications*, volume 10 of *Proceedings of the IMA*, pages 147–186. Springer-Verlag, New York, 1988.

[20] J. M. Harrison. The BIGSTEP approach to flow management in stochastic processing networks. In F. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, pages 57–90. Oxford University Press, 1996.

[21] J. M. Harrison. Heavy traffic analysis of a system with parallel servers: Asymptotic optimality of discrete-review policies. *Ann. Appl. Prob.*, 8:822–848, 1998.

[22] J. M. Harrison. A broader view of brownian networks. 2000. Submitted.

[23] J. M. Harrison. Stochastic networks and activity analysis. In Y. Suhov, editor, *Analytic methods in Applied Probability, In memory of Fridrih Karpelevich*. AMS, Providence, RI, 2002.

[24] J. M. Harrison and J. A. Van Mieghem. Dynamic control of brownian networks: State space collapse and equivalent workload formulations. *Ann. Appl. Prob.*, 7:747–771, 1996.

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

[26] J. M. Harrison and L. M. Wein. Scheduling networks of queues: Heavy traffic analysis of a two-station closed network. *Oper. Res.*, 38(6):1052–1064, Dec. 1990.

[27] F. P. Kelly and C. N. Laws. Dynamic routing in open queueing models: Brownian models, cut constraints and resource pooling. *Queueing Systems*, 13:47–86, 1993.

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

[29] S. Kumar. Two-server closed networks in heavy traffic: Diffusion limits and asymptotic optimality. *Ann. Appl. Prob.*, 10(3):930–961, 2000.

[30] S. Kumar and M. Muthuraman. A numerical method for solving singular Brownian control problems. *Submitted. Short version appeared in Proc. IEEE Conf. on Decision and Control*, 2000.

[31] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*. Springer, New York, 1992.

[32] H. J. Kushner and L. F. Martins. Heavy traffic analysis of a controlled multiclass queueing network via weak convergence methods. *SIAM J. on Control and Optimization*, 34(5):1781–1797, Sept. 1996.

[33] C. Maglaras. Dynamic scheduling in multiclass queueing networks: Stability under discrete-review policies. *Queueing Systems*, 31(3):171–206, 1999.

[34] C. Maglaras. Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality. *Ann. Appl. Prob.*, 10(3):897–929, 2000.

[35] L. F. Martins, S. E. Shreve, and H. M. Soner. Heavy traffic convergence of a controlled multiclass queueing network. *SIAM J. on Control and Optimization*, 34(6):2133–2171, Nov. 1996.

[36] S. P. Meyn. Stability and optimization of queueing networks and their fluid models. In G. G. Yin and Q. Zhang, editors, *Mathematics of Stochastic Manufacturing Systems*, volume 33 of *Lectures in Applied Mathematics*, pages 175–200. American Mathematical Society, 1997.

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

[38] S. P. Meyn. Sequencing and routing in multiclass queueing networks Part II: workload relaxations. 2001. Submitted for publication.

[39] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of optimal queueing network control. *Math. Oper. Res.*, 24(2):293–305, May 1999.

[40] A. N. Rybko and A. L. Stolyar. Ergodicity of stochastic processes describing the operations of open queueing networks. *Problems of Information Transmission*, 28:199–220, 1992.

[41] S. Sethi and G. Sorger. A theory of rolling horizon decision making. *Annals of Operations Research*, 29:387–416, 1991.

[42] A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis: Queues, communications and computing*. Chapman & Hall, 1995.

[43] Y. C. Teh. Dynamic scheduling for queueing networks derived from discrete-review policies. In D. R. McDonald and

S. R. E. Turner, editors, *Analysis of Communication Networks: Call Centres, Traffic and Performance*, volume 28 of *Fields Institute Communications*. American Mathematical Society, 2000.

[44] R. J. Williams. On the approximation of queueing networks in heavy traffic. In F. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, pages 35–56. Oxford University Press, 1996.

[45] R. J. Williams. Diffusion approximations for open multiclass queueing networks: Sufficient conditions involving for state space collapse. *Queueing Systems*, 30:27–88, 1998.

[46] R. J. Williams. An invariance principle for semimartingale reflecting Brownian motion. *Queueing Systems*, 30:5–25, 1998.

[47] P. Yang. Least controls for a class of constrained linear stochastic systems. *Math. Oper. Res.*, 18:275–291, 1993.