COLUMBIA
BUSINESS
SCHOOL

Spring 1999
Professor Costis Maglaras
B9801-33 Stochastic Processing Networks
409 Uris Hall

Part III: Dynamic Control of Fluid Models

Summary. In this note we study dynamic control of fluid models motivated from a class of of stochastic optimal network control problems. In particular, we start by reviewing the basic derivation and properties of fluid models. Then, given a stochastic optimal network control problem, we formulate the associated fluid optimization problem. We describe the basic properties of fluid optimal control problems and the characterization of their solution. We prove stability of the fluid optimal control policy, and subsequently derive further properties regarding the transient behavior of the fluid optimal control solution. We provide a brief overview of numerical optimization algorithms for the solution of these problems. We conclude by considering alternative dynamic control policies for fluid control problems.

The following notation will be useful. Let $\mathbf{R}_+ = [0,\infty)$. $\mathbf{C}_R[0,\infty)$ denotes the space of continuous functions on $\mathbf{R}_+$. $\mathbf{D}_R[0,\infty)$ is the space of right continuous functions on $\mathbf{R}_+$ having left limits on $(0,\infty)$ (RCLL), endowed with the Skorohod topology; see Ethier and Kurtz [EK86, section 3.5]. Given a sequence of functions $\{f_n\}$, where $f_n \in \mathbf{D}_R[0,\infty)$ for each n and a function $f \in \mathbf{C}_R[0,\infty)$, $f_n \to f$ in the Skorohod topology if and only if $f_n \to f$ uniformly on compact sets (u.o.c.). That is,

$$\sup_{0\le s\le t} |f_n(s) - f(s)| \to 0, \quad \text{as} \quad n \to \infty.$$

The vector of ones of appropriate dimension will be denoted by $\mathbf{1}$, $\lfloor x \rfloor$ will be the integer part of x rounded down, $x \wedge y = \min(x,y)$, $x \vee y = \max(x,y)$, $x^+ = \max(0,x)$, and finally all vector equalities or inequalities should be interpreted componentwise. The transpose of a matrix P will be denoted P', $\mathbf{diag}\{x_1,\ldots,x_n\}$ will denote the $n \times n$ diagonal matrix with diagonal elements $(x_1,\ldots,x_n)$.

III.1 Fluid models: review and properties

We use the following convention: a generic vector of queue lengths (or fluid levels) will be denoted by z, and the initial queue length configuration will be denoted by q.

Recall the equation that captures the system dynamics:

$$Z(t) = q + E^x(t) + \sum_{k=1}^{K} \Phi^k(D_k^x(t)) - D^x(t), \tag{III.1}$$

where x is the initial condition (of the Markov state descriptor that depends on q and other auxiliary quantities) and $D_k(t)$ is the number of class k departures up to time t defined by

$$D_k^x(t) = \max\{i : v_k(1) + \cdots + v_k(i) \le T_k^x(t)\}.$$

Consider a sequence of initial conditions $\{x_n\} \subset \mathbf{X}$ such that $|x_n| \to \infty$ as $n \to \infty$ and for any real valued process $\{f(t), t \geq 0\}$ define its fluid scaled counterpart by

$$\bar{f}^n(t) = \frac{1}{|x_n|} f^{x_n}(|x_n|t). \tag{III.2}$$

In the sequel, the overbar notation will signify fluid scaled quantities and appropriate superscripts will be used to signify the scaled processes corresponding to some initial condition along the sequence $\{x_n\}$.

A functional strong law of large numbers (FSLLN) can be derived for the fluid-scaled processes defined above. The following is a restatement of Proposition 1.2.4 [Dai98].

Lemma III.1.1 *Let $\{x_n\} \subset \mathbf{X}$ such that $|x_n| \to \infty$ as $n \to \infty$. Assuming that*

$$\lim_{n\to\infty} \bar{R}_a^n(0) = \frac{1}{|x_n|} R_a^n(0) = \bar{R}_a \quad \text{and} \quad \lim_{n\to\infty} \bar{R}_s^n(0) = \frac{1}{|x_n|} R_s^n(0) = \bar{R}_s,$$

as $n \to \infty$, almost surely

$$\frac{1}{|x_n|} \Phi^k(\lfloor |x_n| t \rfloor) \to P_k' t, \quad u.o.c. \tag{III.3}$$

$$\frac{1}{|x_n|} E_k^n(\lfloor |x_n| t \rfloor) \to \alpha_k (t - \bar{R}_a)^+, \quad u.o.c. \tag{III.4}$$

$$\frac{1}{|x_n|} D_k^n(\lfloor |x_n| t \rfloor) \to \mu_k (t - \bar{R}_s)^+, \quad u.o.c.. \tag{III.5}$$

The limit processes in (III.3)-(III.5) are deterministic and continuous. Continuity follows from the fact that the jump size of the scaled processes is decreasing as $1/|x_n|$, which yields continuous limit trajectories. The deterministic nature of these limits is a direct consequence of the FSLLN scaling. Note that the limits $\bar{R}_a, \bar{R}_s$ depend on the sequence $\{x_n\}$.

Applying the scaling of (III.2) to (III.1) and using (III.3)-(III.5) we get the following result which is an elaboration of Theorem 2.3.1 in [Dai98].

Theorem III.1.1 *[Dai98, Theorem 4.1] For almost all sample paths ω and any sequence of initial conditions $\{x_n\} \subset \mathbf{X}$ such that $|x_n| \to \infty$ as $n \to \infty$, there is subsequence $\{x_{n_j}(\omega)\}$ with $|x_{n_j}(\omega)| \to \infty$ such that*

$$(\bar{Z}^{x_{n_j}}(0,\omega), \bar{R}_a^{x_{n_j}}(0,\omega), \bar{R}_s^{x_{n_j}}(0,\omega) \to (z(0,\omega), \bar{R}_a(\omega), \bar{R}_s(\omega)) \tag{III.6}$$

$$(\bar{Z}^{x_{n_j}}(t,\omega),\ \bar{T}^{x_{n_j}}(t,\omega)) \to (z(t,\omega),\ \bar{T}(t,\omega)) \quad u.o.c.. \tag{III.7}$$

Furthermore, $(z(\cdot,\omega), \bar{T}(\cdot,\omega))$ satisfies the following set of equations

$$z(t) = z(0) + \alpha(t\mathbf{1} - \bar{R}_a)^+ - (I - P')M^{-1}(\bar{T}(t) - \bar{R}_s)^+, \tag{III.8}$$

$$z(t) \geq 0 \quad for\ t \geq 0, \tag{III.9}$$

$$\bar{I}(t) = \mathbf{1}t - C\bar{T}(t),\ \bar{T}(0) = 0, \tag{III.10}$$

$$\bar{T}(t),\ \bar{I}(t) \quad are\ non\text{-}decreasing\ for\ t \geq 0, \tag{III.11}$$

together with some additional conditions on $(z(\cdot,\omega), \bar{T}(\cdot,\omega))$ *that are specific to the scheduling policy employed.*

That is, the fluid limits depend on ω and on the converging subsequence $\{x_{n_j}\}$ as well (through $\bar{R}_s, \bar{R}_a$). They are neither deterministic nor unique, but their dynamics are captured by these deterministic and continuous equations of evolution (III.8)-(III.11). Hereafter. whenever possible the dependence on ω will be suppressed from the notation.

The above set of equations will be referred to as the *delayed fluid model* associated with a multiclass queueing network under a specified scheduling policy. Moreover, we will say that $(z, \bar{T}) \in FM$ -or equivalently that it is a *fluid solution*- if this pair of state and input trajectories satisfies equations (III.8)-(III.11). It is immediate from (III.8)-(III.11) that the limit processes $(z, \bar{T})$ are Lipschitz continuous. Hence, it follows that that they have a time derivative almost everywhere. A path $q(\cdot)$ is called regular at t if it is differentiable at t and its derivative at time t will be denoted by $\dot{q}(t)$. Let $v(t)$ denote the instantaneous *fluid* allocation vector at time t. The cumulative allocation process can be rewritten as

$$\bar{T}(t) = \int_0^t v(s)ds, \ t \geq 0.$$

Restricting attention to the case where $\bar{R}_a = \bar{R}_s = 0$ and using the a.e. differentiability of the limit processes, for almost all times $t \geq 0$ the fluid limit model can be expressed as a linear dynamical system with polytopic constraints in $v(t)$ of the following form:

$$\dot{z}(t) = \alpha - Rv(t), \quad z(0) = q, \tag{III.12}$$

$$z(t) \geq 0, \quad Cv(t) \leq \mathbf{1}, \quad v(t) \geq 0 \ \text{ for } t \geq 0, \tag{III.13}$$

together with some policy specific conditions. The fluid limit model in (III.12)-(III.13) is called *undelayed.* This is the only case studied in the sequel. Once again, the dependence of the fluid limits on the sample path ω has been suppressed. (In the literature the term undelayed refers to the set of integral equations (III.8)-(III.11) for the case where $\bar{R}_a = \bar{R}_s = 0$.) Following our earlier notation we will say that that $(z, v) \in FM$ -or equivalently that it is a *fluid solution*- if this pair of state and input trajectories satisfy equations (III.12)-(III.13). Undelayed limits can be obtained if one restricts attention to exponential interarrival and service time processes, or in the case of general distributions, if one lets $|x_n| \to \infty$ while keeping $R_a^n(0)$ and $R_s^n(0)$ bounded.

III.2 Formulation of the control problem

III.2.1 Queueing network control problem

The canonical class of network control problems we address is as follows. Let $g : \mathbf{R}_+^K \to \mathbf{R}_+$ be a $\mathbf{C}^2$ convex cost rate function such that, for some constants $\underline{b}, \underline{c}, \bar{b}, \bar{c} > 0$ such that $\underline{b} \leq \bar{b}$, $\underline{c} \leq \bar{c}$, and

$$\underline{b}|x|^{\underline{c}} \leq g(x) \leq \bar{b}|x|^{\bar{c}}. \tag{III.14}$$

Note that (III.14) implies that $g(x) = 0 \Leftrightarrow x = 0$. Given the cost rate function g, the following stochastic network control problem is considered: choose an allocation process $T(t)$, or equivalently, an admissible policy π, in order to minimize

$$J_T^\pi(q) = \mathbf{E}_q^\pi \int_0^T g(Z(t))dt, \tag{III.15}$$

where $\mathbf{E}_q^\pi$ denotes the expectation operator with respect to the probability measure $\mathbf{P}_q^\pi$ defined by any admissible policy π and initial condition q. The use of T with no time argument will denote a time horizon and should not be confused with the cumulative allocation $T(t)$.

In this problem formulation attention is restricted to finite horizon network control problems. It is convenient to think of T as being long but finite. The objective in these problems remains finite starting from an arbitrary (but finite) initial condition independent of the traffic intensity (or load) at each station, and in particular, this problem remains meaningful even when $\rho \not\leq \mathbf{1}$ where, for example, long run averages will not exist. This will allow for an easy extension to the *heavy-traffic* regime, where $\rho \to \mathbf{1}$. This case will be addressed later on.

Complexity. Let $\mathcal{P}$ denote the class of problems that can be solved in polynomial running time. Let $\mathcal{NP}$ denote the class of problems for which given a polynomial size certificate, we can check in polynomial time whether this certificate is valid. A useful attribute of complexity is *hardness.* A problem L is $\mathcal{NP}$-hard if every problem in class $\mathcal{NP}$ can be reduced in polynomial time to problem L; L need not belong to class $\mathcal{NP}$ itself – if it does then it is called $\mathcal{NP}$-complete. Let $\mathcal{EXPTIME}$ be the class of problems that can be solved in exponential running time. Clearly, $\mathcal{P} \subseteq \mathcal{NP}$ and $\mathcal{P} \neq \mathcal{EXPTIME}$.

The intractability of the class of problems considered here was formally established in a very strong sense by Papadimitriou and Tsitsiklis [PT96], where they showed that the multiclass network problem is $\mathcal{EXPTIME}$-hard. This immediately implies that this problem is of exponential complexity. (This is true even if one could show that $\mathcal{P} = \mathcal{NP}$, which is by itself quite unlikely).

In light of these negative complexity results, the approach we will follow here is based on approximating (or replacing) the stochastic network by its fluid analog -this is a model with deterministic and continuous dynamics-, solving an associated fluid optimal control problem, and then using the derived fluid control policy in order to define an implementable policy in the stochastic network. This procedure is summarized below:

1. Consider a dynamic control problem for the original stochastic network;
2. Form fluid analog of stochastic network and solve the associated fluid optimal control problem;
3. Translate/implement the optimal fluid control in original stochastic network;
4. Consider fluid limit of stochastic network under implemented policy;
5. Verify fluid-scale asymptotic optimality and stability.

Stages 1 to 3 are clear. Stages 4 and 5 describe a criterion for performance analysis under the implemented policy that is consistent with the model approximation adopted at stage 2, in the following sense: the implementation is tested for asymptotic optimality in the limiting regime where the model approximation is valid. This criterion is referred to as *fluid-scale asymptotic optimality* (FSAO). We will return to this flowchart later on. Now we will focus on stage 2.

III.2.2 Fluid optimal control problem

We start by formulating the associated fluid model. Discrete jobs moving stochastically through different queues are replaced by continuous fluids flowing through different buffers, and system evolution is observed starting from any initial state. The deterministic rates at which the different fluids flow through the system are given by the average rates of corresponding stochastic quantities. The fluid optimization problem associated with (III.15) is defined by

$$\bar{V}^g(q) = \min_{v(\cdot)} \left\{ \int_0^T g(z(t))dt : z(0) = q \text{ and } (z, v) \in FM \right\}. \tag{III.16}$$

$\bar{V}^g(q)$ denotes the value function of the fluid optimization problem starting from the initial condition q, the superscript g denotes the dependence on the cost rate function, and T is the same time horizon that appears in the performance index in (III.15). The problem in (III.16) is one of *transient optimization* or *transient recovery* starting from a large initial backlog. One should think of T as being long enough so that starting from any appropriately normalized initial condition z, the transient behavior of the optimal solution will have settled and the queue length vector will have reached its final state of least achievable cost without being affected by the finite horizon T; more on this later. Specifically, optimization in the fluid model gives information about the path that the state will follow until it reaches a "final" state; for $\rho < \mathbf{1}$ this state is at the origin, for $\rho = \mathbf{1}$ this is the state of minimum achievable cost given the initial condition, and if $\rho > \mathbf{1}$ this state will correspond to the asymptote of minimum cost accumulation, along which the fluid trajectory will blow up as t increases.

III.2.3 An example

We will analyze the Rybko-Stolyar network shown in Figure 1. For illustrative purposes we shall consider the following specific numerical data:

$$\alpha_1 = \alpha_3 = 1, \quad \mu_1 = \mu_3 = 6 \text{ and } \quad \mu_2 = \mu_4 = 1.5. \tag{III.17}$$

Control capability in this network is with regard to sequencing decisions between classes 1 and 4 at server 1 and classes 2 and 3 at server 2. Note that the two job classes waiting to be processed at each server differ in their service requirements and routes through the network. Now suppose we wish to

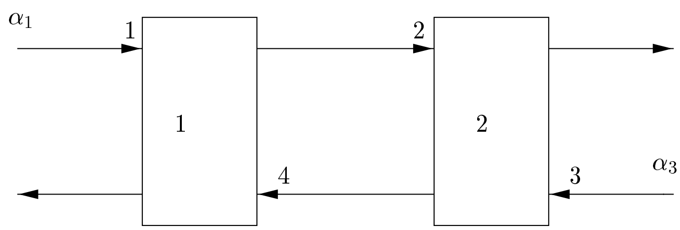

Figure 1: The *Rybko-Stolyar* network

find a scheduling policy π that minimizes

$$J_\pi(T) = \mathbf{E}^\pi \int_0^T \sum_{k=1}^4 Z_k(t)dt, \tag{III.18}$$

where $Z_k(t)$ is the class k queue length at time t, and $\mathbf{E}^\pi$ denotes the expectation operator with respect to the probability measure $\mathbf{P}^\pi$ defined by any admissible policy π.

Specifically, for the Rybko-Stolyar network the fluid model equations are as follows. Denoting by $v_k(t)$ the instantaneous fraction of effort devoted to serving class k jobs at time t by the associated server, and by $z_k(t)$ the amount of fluid in buffer k at time t, and defining vector functions $v(t)$ and $z(t)$ in the obvious way, one has

$$\dot{z}(t) = \alpha - Rv(t), \quad z(0) = q, \tag{III.19}$$

$$v(t) \geq 0, \quad v_1(t) + v_4(t) \leq 1, \quad v_2(t) + v_3(t) \leq 1, \quad z(t) \geq 0, \tag{III.20}$$

where

$$\alpha = \begin{bmatrix} \alpha_1 \\ 0 \\ \alpha_3 \\ 0 \end{bmatrix}, \quad R = \begin{bmatrix} \mu_1 & 0 & 0 & 0 \\ -\mu_1 & \mu_2 & 0 & 0 \\ 0 & 0 & \mu_3 & 0 \\ 0 & 0 & -\mu_3 & \mu_4 \end{bmatrix}.$$

The associated fluid optimal control problem will be to choose a control $v(\cdot)$ for this fluid model that minimizes

$$\bar{J}(z,T) = \int_0^T \sum_{k=1}^4 z_k(t)dt, \tag{III.21}$$

for some fixed $T > 0$. The corresponding value function will be denoted by $\bar{V}(z,T)$.

It can be shown that the optimal control for the fluid model can be characterized as *Last-Buffer-First-Served (LBFS)* with server splitting whenever an exiting class is emptied at the other server. That is, each server has responsibility for one incoming buffer and one exit buffer; the exit buffer is given priority unless the other server's exit buffer is empty, in which case server splitting occurs. For an explanation of the latter situation, let us focus on the behavior of server 1 when buffer 2 (the exit buffer for server 2) is empty and buffer 1 is non-empty. In that circumstance, given the data in

$z^*(t)$	$v^*(t)$
$(*,+,*,+)$	$(0,1,0,1)$
$(+,0,0,0)$	$(1-\alpha_3/\mu_4, 1-\alpha_3/\mu_3, \alpha_3/\mu_3, \alpha_3/\mu_4)$
$(+,0,*,+)$	$(\mu_2/\mu_1, 1, 0, 1-\mu_2/\mu_1)$
$(*,+,+,0)$	$(0, 1-\mu_4/\mu_3, \mu_4/\mu_3, 1)$
$(0,0,+,+)$	$(\alpha_1/\mu_1, \alpha_1/\mu_2, 1-\alpha_1/\mu_2, 1-\alpha_1/\mu_1)$
$(+,+,0,0)$	$(1-\alpha_3/\mu_4, 1-\alpha_3/\mu_3, \alpha_3/\mu_3, \alpha_3/\mu_4)$
$(0,0,0,+)$	$(\alpha_1/\mu_1, \alpha_1/\mu_2, \alpha_3/\mu_3, 1-\alpha_1/\mu_1)$
$(0,+,0,0)$	$(\alpha_1/\mu_1, 1-\alpha_3/\mu_3, \alpha_3/\mu_3, \alpha_3/\mu_4)$
$(+,0,+,0)$	$(\mu_2, \mu_1, \mu_4, \mu_3)/(\mu_1+\mu_2)$
$(0,0,+,0)$	$(\alpha_1/\mu_1, \alpha_1/\mu_2, 1-\alpha_1/\mu_2, 1-\alpha_1/\mu_1)$
$(0,0,0,0)$	$(\alpha_1/\mu_1, \alpha_1/\mu_2, \alpha_3/\mu_3, \alpha_3/\mu_4)$

Table 1: Optimal control for the fluid model associated with the Rybko-Stolyar network

(III.17), server 1 devotes 25% of its effort to buffer 1 (its own incoming buffer) so that server 2 can remain fully occupied with class 2 jobs, and devotes the other 75% of its effort to draining buffer 4 (its own exit buffer). This policy is myopic in the sense that it removes fluid from the system at the fastest possible instantaneous rate, regardless of future considerations, and it is optimal regardless of the horizon length T. This yields the following alternative characterization:

$$v(t) \;\in\; \text{argmin} \;\left\{\mathbf{1}'\dot{z}(t) : v \geq 0,\; v_1+v_4 \leq 1,\; v_2+v_3 \leq 1,\; z(t) \geq 0\right\}. \qquad \text{(III.22)}$$

In more detail, the optimal control is the one described in Table 1, where a "+" signifies positive buffer content, a "0" signifies an empty buffer, a "$*$" signifies arbitrary buffer content for the state vector $z(t)$, and $v(t)$ is the optimal instantaneous allocation vector. Finally, an example of optimal fluid trajectories starting from the initial condition $q = [1, 0, 0.5, 1]$ is depicted in Figure 2.

III.3 Characterization of the Optimal Control

1. *Existence.* First, note that the control $v(t) = R^{-1}\alpha$ for $0 \leq t \leq T$ is feasible and moreover, it provides the following upper bound on the value function: $0 \leq \bar{V}^g(q) \leq Tg(q)$. Moreover, given the properties of the cost rate function g, it follows that the objective in (III.16) is continuous in z and thus, the resulting optimization problem is of the form of minimization of a continuous function in a compact set. Existence of an optimal solution follows from Weiertrass Theorem.

2. *Continuity & differentiability.* Given a compact set of initial conditions $\{q : |q| \leq B\}$, and using

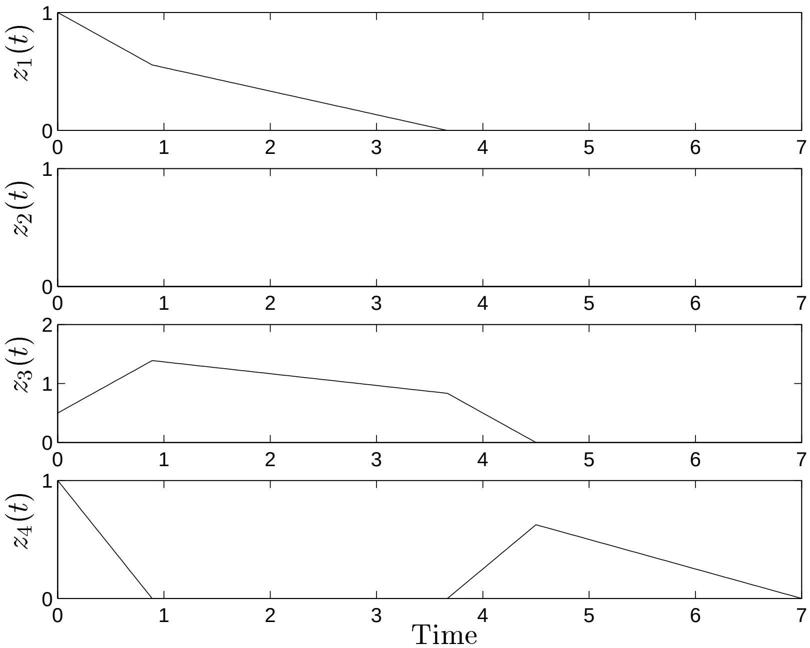

Figure 2: Optimal state trajectories in the fluid model for $q = [1, 0, 0.5, 1]$

the Lipschitz continuity of the fluid trajectories $z(\cdot)$, we can bound above the set of feasible fluid solutions of (III.16) by $BT\kappa$, where κ is an appropriate growth constant determined by α and R. Then $G = \sup\{\nabla g(z)'(\alpha - Rv) : v \in \mathcal{V}(z),\ |z| \le BT\kappa\}$, where $\mathcal{V}(z) = \{v : v \ge 0,\ Cv \le \mathbf{1},\ (Rv)_k \le \alpha_k$ for all k such that $z_k = 0\}$ is the set of admissible controls when the state is z. Note that the set of admissible controls is non-empty, since $R^{-1}\alpha \in \mathcal{V}(z)$ for all $z \ge 0$. Using this definition, G is a Lipschitz constant for g and consequently for the objective of (III.16) as well. This implies almost everywhere differentiability of the value function $\bar{V}^g$ for the set of initial conditions $|q| \le B$. (This is related to the idea of *locally Lipschitz* functions described in Dai's notes [Dai98].

A generalized derivative can be defined (by choosing an appropriate subgradient) at points where $\bar{V}^g$ is not differentiable. Since g is $\mathbf{C}^2$ it follows that the generalized gradient of $\bar{V}^g$, denoted $\nabla\bar{V}^g$, will also be a.e. continuous.

3. *Smooth approximation.* Hereafter, we proceed under the stronger assumption that $\nabla\bar{V}^g$ is in fact everywhere continuous; this assumption is not restrictive, since one could always construct a smooth approximation of any optimal fluid trajectory that is arbitrarily close to it, and proceed by analyzing this smooth approximation (that has a continuous gradient) thereafter.

4. *Characterization of the optimal control.* The optimal instantaneous allocation is a state feedback law that can be characterized by a direct application of dynamic programming principles. The

following is an informal derivation of the optimality conditions for the solution of this problem in terms of the *cost-to-go* function of the optimal control problem, denoted $\bar{V}^g(z(t),t)$. Note that the cost-to-go is a function of the current state and the time t; this is a consequence of the finite control horizon T. In this notation, $\bar{V}^g(q)$ should read $\bar{V}^g(q,0)$.

$$\begin{aligned}\bar{V}^g(z(t),t) &= g(z(t))\delta t + \min_{v\in\mathcal{V}(z(t))} \bar{V}^g(z(t+\delta t), t+\delta t) + o(\delta t) &\text{(III.23)}\\ &= g(z(t))\delta t + \bar{V}^g(z(t),t) + \frac{d}{dt}\bar{V}^g(z(t),t)\delta t + \\ &\qquad + \min_{v\in\mathcal{V}(z(t))} \nabla\bar{V}^g(z(t),t)'(\alpha - Rv)\delta t + o(\delta t). &\text{(III.24)}\end{aligned}$$

The optimal control $v(t)$ is computed as the solution of the following linear program

$$v(t) = \operatorname*{argmin}_{v\in\mathcal{V}(z(t))} \nabla\bar{V}^g(z(t),t)'(\alpha - Rv) = \operatorname*{argmax}_{v\in\mathcal{V}(z(t))} \nabla\bar{V}^g(z(t),t)'Rv. \qquad \text{(III.25)}$$

III.4 Stability Analysis of the Fluid Optimal Control Policy

Before we proceed to interpret this equation and review the structural properties of the solution, we prove stability of the optimally controlled fluid model. In passing we obtain an upper bound on the control horizon T required in order to get a stationary solution starting from any initial condition with $|q| \leq 1$. This will allow us to avoid dealing with the non-stationary nature of the control policy described above.

Recall the definition of stability: the fluid model associated with a scheduling policy is *stable* if there exists a time $T > 0$ such that for any solution $z(\cdot)$ of the fluid model equations with $|z(0)| = 1$, $z(t) = 0$, for $t \geq T$. The next proposition shows that the optimally controlled fluid model is stable provided that $\rho < \mathbf{1}$.

Proposition III.4.1 *Assuming that $\rho < \mathbf{1}$, there exists a constant T_g that depends on the cost rate function $g(\cdot)$, such that for any $T > T_g$ the fluid model (III.12), (III.13) and (III.25) is stable.*

Proof. . Given an initial condition q, an input control $\hat{v}(t)$ will be constructed that will linearly translate the state, starting from q back to the origin. (In the sequel, all quantities related to this construction will be denoted by a "hat".) Let $\hat{T}(t)$ be the total allocation process associated to the instantaneous input control $\hat{v}(t)$ and let t^* be the time that the fluid network will empty under this control. Rewrite equation (III.12) in the form

$$\hat{z}(t) = q + R\hat{Y}(t), \quad \text{where} \quad \hat{Y}(t) = R^{-1}\alpha t - \hat{T}(t).$$

In this case, $\hat{Y}(t^*) = -R^{-1}q$ and thus $\hat{T}(t^*) = R^{-1}\alpha t^* + R^{-1}q$. Linear translation from q back to the origin implies that $\hat{v}(t) = \hat{T}(t^*)/t^*$ for all $t \leq t^*$. By the definition of R it follows that $R^{-1} =$

$M(I-P')^{-1} = M(I+P+P^2+\cdots)$, which is elementwise non-negative and thus the instantaneous control defined above satisfies the constraint $\hat{v}(t) \geq 0$. Next, the capacity constraints imply that

$$C\hat{v}(t) \leq \mathbf{1} \;\;\Rightarrow\;\; \rho + \frac{CR^{-1}q}{t^*} \leq \mathbf{1} \;\;\Rightarrow\;\; \frac{CR^{-1}q}{t^*} \leq \mathbf{1} - \rho \;\;\Rightarrow\;\; t^* \geq \max_i \frac{(CR^{-1}q)_i}{1-\rho_i}.$$

Under the input control $\hat{v}(t)$, the resulting state trajectory is described by

$$\hat{z}(t) = q\left(1 - \frac{t}{t^*}\right) \quad \text{for} \quad t \leq t^* \quad \text{and} \quad \hat{z}(t) = 0 \quad \text{for} \quad t > t^*, \tag{III.26}$$

which clearly satisfies the state positivity constraint $\hat{z}(t) \geq 0$.

Let $\hat{V}^g(q)$ be the total cost accrued in the fluid model under the input $\hat{v}(t)$, which can be computed from the expression

$$\hat{V}^g(q) = \int_0^{t^*} g\left(q(1-t/t^*)\right) dt.$$

This is also an upper bound on the value function $\bar{V}^g(q)$. (This upper bound is valid even if $t^* \geq T$.)

Given the control horizon T in (III.16), $\min_{t \leq T} g(z(t)) \leq \hat{V}_g(q)/T$, where $z(t)$ is the state trajectory under the policy defined by (III.25) that achieves the minimum draining cost $\bar{V}^g(q)$. Let τ be the time that this minimum is attained. Given the properties of $g(\cdot)$, we have that $g(z(\tau)) \geq \underline{b}|z(\tau)|^{\underline{c}}$, which implies that $|z(\tau)|^{\underline{c}} \leq \hat{V}_g(q)/(T\underline{b})$. Next we choose the control horizon T long enough such that for any $0 < \gamma < 1$ we have that $|z(\tau)| \leq \gamma$, independent of the initial condition q. Let

$$\zeta = \max_{|q|=1} \; \max_i \frac{(CR^{-1}q)_i}{1-\rho_i} \quad \text{and} \quad \delta = \max_{|q|\leq 1} g(q) = \bar{b}. \tag{III.27}$$

For any initial condition such that $|q| = 1$, $\hat{V}_g(q) \leq \delta t^* \leq \bar{b}\zeta$. Let $T_g = \bar{b}\zeta/(\gamma^{\underline{c}}\underline{b})$. Then, for any $T > T_g$, there exists a time $0 < \tau \leq T$ such that $|z(\tau)| \leq \gamma$.

The remainder of this proof imitates the arguments in Theorem 6.1 of Stolyar [Sto95]. For $m = 1, 2, \ldots$, let $\tau_m = \min\{t > 0 : |z(t)| \leq \gamma^m,\ |z(0)| = \gamma^{m-1}\}$. Modifying (III.27), we can define $\zeta^m = \zeta\gamma^{m-1}$ and $\delta^m = \bar{b}(\gamma^{m-1})^{\bar{c}}$. It follows that

$$\tau_m \leq T_g(\gamma^{\bar{c}+1-\underline{c}})^{m-1}.$$

Clearly, $\sum_m \tau_m \leq T_g/(1-\gamma^{\bar{c}+1-\underline{c}}) \triangleq T_0$. Continuity of $|z(t)|$ in t, implies that $\lim_m |z(\sum_m \tau_m)| = 0$, and therefore, that $\sup\{t > 0 : |z(t)| = 0,\ |z(0)| = 1\} \leq \sum_m \tau_m \leq T_g/(1-\gamma^{\bar{c}+1-\underline{c}})$. By observing that the fluid trajectories under the optimal fluid control policy will remain empty once they drain for the first time we complete the proof. $\diamond$

While this result is hardly surprising due to the optimality implied by (III.25), its derivation provided an upper bound for the time taken to drain the fluid model starting from any initial condition with $|q| = 1$ under the infinite horizon optimal control solution; (III.16) is still well-posed when $T = \infty$, (III.25) is a characterization of the optimal control, and Proposition III.4.1 is still valid, but, in fact, the optimal control $v(t)$ no longer depends on t.

Intuitively, the geometric convergence observed in the proof is due to the Lipschitz paths of the optimal control; that is, even is $|z(t)|$ is small, the rate of decay is still linear and this yields the desired result. This should be contrasted with the case where, for example, $\dot{z}(t) = Az(t)$ where A is s stable matrix, i.e., all eigenvalues of A have negative real part. There, the drift decreases as $|z(t)|$ gets small, which results in the exponential decay $z(t) = e^{At}q$. In this case, there is no time T such that $z(t) = 0$ for $t \geq T$; i.e., this system is not stable according to our definition. For a person familiar with system theory our original definition of stability might have appeared to be too strict; This person would have asked: "Isn't exponential decay to 0 sufficient? It must be!" In closer look, however, we see that our definition has the correct form, since it picks up the essential dynamics of our dynamical systems that are different than the classical models in linear system theory.

Corollary III.4.1 *Let $z^*(\cdot)$ denote any optimal trajectory for (III.16) with $T = \infty$, starting from any initial condition with $|q| \leq 1$. For any $\gamma \in (0,1)$, $z^*(t) = 0$ for $t \geq T_0$, where*

$$T_0 \triangleq \frac{\bar{b}\ \zeta}{\gamma^{\underline{c}}\ \underline{b}}\ \frac{1}{1 - \gamma^{\bar{c}+1-\underline{c}}}. \tag{III.28}$$

Minimizing over γ one gets the smallest upper bound T_0 at

$$\gamma^* = \left(\frac{\underline{c}}{\bar{c}+1}\right)^{1/(\bar{c}+1-\underline{c})}.$$

The estimate of T_0 is important in numerical optimization routines.

If we set $T \geq T_0$, (III.25) reduces to the stationary solution of the infinite horizon fluid control problem

$$v(t) \ = \ \operatorname*{argmin}_{v \in \mathcal{V}(z(t))} \ \nabla \bar{V}^g(z(t))'(\alpha - Rv) \ = \ \operatorname*{argmax}_{v \in \mathcal{V}(z(t))} \ \nabla \bar{V}^g(z(t))' Rv. \tag{III.29}$$

A similar condition can be derived when $\rho = \mathbf{1}$. Hereafter, it is assumed that (III.28) is satisfied.

III.5 Structural Properties of the Optimal Control

1. *Bang-bang.* The following fact explains some of the structural properties of the solution of the fluid optimal control solution. They are consequences of the deterministic and continuous dynamics of fluid models and the fact that we have bounded controls ($v(t) \in [0,1]$). First, we need some definitions.

   - The sequence of time epochs $P = \{t_0, t_1, \ldots, t_p\}$ is a partition of $[0,T]$ if $0 = t_0 \leq t_1 \leq \cdots \leq t_p = T$. These points will be referred to t_i as breakpoints.

- A function $f(t)$ is piecewise constant (linear) on a partition P, if it is constant (linear) on $[t_{i-1}, t_i)$ for $i = 1, \ldots, p$. A function $f(t)$ is piecewise constant (linear) on $[0, T]$ if it is piecewise constant (linear) with some partition of $[0, T]$.

Fact III.5.1 *Let $(z_q^*(\cdot), v_q^*(\cdot))$ denote the optimal fluid and control trajectory associated with (III.16) (where q is the initial condition).*

(a) $v_q^*(\cdot)$ is piecewise constant for some partition P of finitely many breakpoints.

(b) $z_q^*(\cdot)$ is continuous and piecewise linear on $[0, T]$.

(c) If $z_q^*(t) > 0$, then $v_q^*(t)_k$ equal either 0 or 1 (this is the bang-bang nature).

In fact, z_q^* will be Lipschitz continuous for the appropriate growth constant that depends on α, R. These points can be observed in the optimal fluid trajectories depicted in Figure 2.

2. *Interpretation of the optimal control.* Define the vector valued function

$$r^g(z(t)) = R' \nabla \bar{V}^g(z(t)).$$

Under the standing assumptions, $r^g(\cdot)$ is a continuous function of its argument. This is a dynamic index rule, or reward function, that defines the optimal policy for the fluid model. It associates with each queue length vector $z(t)$ a corresponding K-vector $r^g(z(t))$ of reward rates for effort devoted to processing the various classes of jobs. Under this interpretation, the optimal control at any point in time simply maximizes total reward over the set of admissible controls. Overloading notation, this dynamic reward function will also be referred to as the optimal policy for the fluid control problem described in (III.16).

For example, consider the case where $z(t) > 0$. In this case, $\mathcal{V}(z(t)) = \{v : v \geq 0, Cv \leq \mathbf{1}\}$, and the optimal instantaneous allocation computed from (III.29) will be to allocate all the effort of each server j to class $k \in \mathcal{C}(j)$ that has the highest instantaneous reward at this station ($r_k^g(z(t)) > r_l^g(z(t))$ for all $l \in \mathcal{C}(j), l \neq k$). That is, serve the class with the highest reward rate at each station.

3. *Switching surfaces.* Pursuing further the last example, the optimal instantaneous allocation will only change if the relative priorities between the different classes change at any server change, or if one of the classes gets depleted. Breakpoints correspond to time instances where the reward rates corresponding to the optimal trajectory switch priorities in the fashion just described.

Thus, the state space $\mathbf{R}_+^K$ will be divided into regions according to the relative orderings induced by $r^g(\cdot)$. These regions are separated by *switching surfaces*. Identifying these switching planes is, in general, hard. The following fact leads to a clear mental picture:

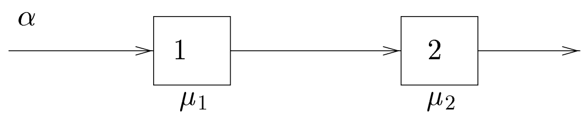

Figure 3: Tandem line - 2 stations

Fact III.5.2 *If the cost rate function g is linear, then the switching surfaces are of the form $\{z : h_i' z = 0, z \geq 0\}$; i.e., they are halfspaces through the origin restricted in the positive orthant.*

We now describe a simple example that will help highlight this fact. Finally, note that non-linear cost rate functions would lead to more complex descriptions for the switching surfaces.

4. *An example.* Consider the example shown in Figure 3. Class 1 jobs arrive with a rate α and after they receive service at station 1 they become class 2 jobs waiting to be served at station 2. We consider the optimal control of this fluid model where the associated cost functional is a linear holding cost: $g(z(t)) = c_1 z_1(t) + c_2 z_2(t)$. The only scheduling decision to be made is whether or not to idle the upstream server.

   Depending on the choice of network parameters the optimal policy will be different. Optimal allocations (for the case $z(t) > 0$) are summarized in Table 2.

	Parameter regime	$v^*(t)$
Case 1	$c_1 \geq c_2$	$(1,1)$
Case 2	$c_1 < c_2$ $\mu_1 \geq \mu_2$	$(0,1)$
Case 3	$c_1 < c_2$ $\mu_1 < \mu_2$ $\frac{z_1(t)}{z_2(t)} < \frac{c_2}{c_1}\frac{\mu_1 - \alpha}{\mu_2 - \mu_1}$	$(0,1)$
Case 4	$c_1 < c_2$ $\mu_1 < \mu_2$ $\frac{z_1(t)}{z_2(t)} \geq \frac{c_2}{c_1}\frac{\mu_1 - \alpha}{\mu_2 - \mu_1}$	$(1,1)$

Table 2: Fluid optimal control for 2-station tandem network

Let's go through these cases.

- In case 1, the holding cost for jobs at server 1 is higher than that at server 2, and thus it is optimal to hold our inventory in the downstream server. Therefore, we never intentionaly

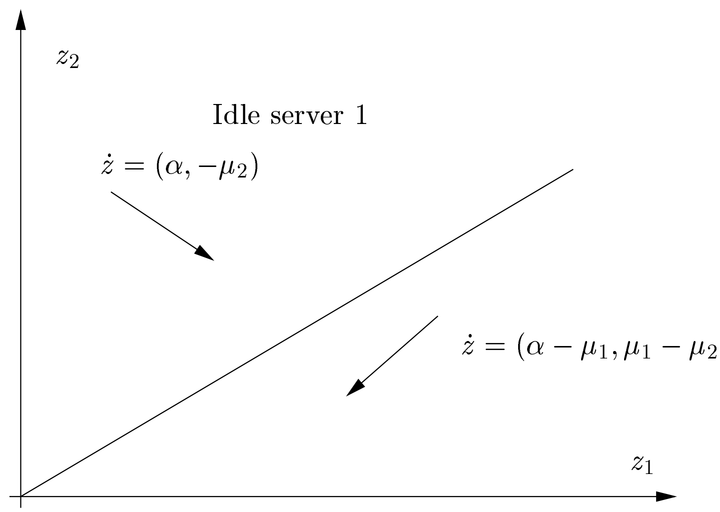

Figure 4: Two stations in tandem: switching line at $\dfrac{z_1(t)}{z_2(t)} = \dfrac{c_2}{c_1}\dfrac{\mu_1 - \alpha}{\mu_2 - \mu_1}$

idle server 1; whenever there is work at station 1 we process it at full capacity. Server 2 continues working at full capacity provided that $z_2(t) > 0$.

- In case 2, it is more expensive to hold inventory in the downstream server and moreover, the upstream server can process jobs faster than the downstream one. Hence, the optimal policy would hold all inventory at server 1, and then process at server 1 at such a rate in order to keep server 2 working at full capacity and its buffer empty; that is, $\mu_1 v_1^* = \mu_2$. If $z(t) > 0$, then server 2 works at full capacity to deplete its queue, whereas server 1 idles until z_2 empties and we get into the scenario just described. This example illustrates that idling could be desirable in the solution of the fluid optimal control problem.
- Cases 3 and 4 are more subtle. Since $\mu_2 > \mu_1$, if queue was depleted, server 1 would not have enough processing capacity to keep server 2 busy and that would result in unwanted idleness in server 2. Therefore, idling server 1 until queue 2 gets depleted as in case 2 is suboptimal. In order to avoid unwanted idleness at server 2 we need to start processing at server 1 before we deplete queue 2. However, since the holding cost at queue 2 is higher than that in queue 1, we will not process at server 1 unless queue 2 starts getting small enough that eventually that we would incur idleness that will prevent us from optimal cost draining. This tradeoff between holding costs at queue 2 or unwanted idleness at the corresponding server leads to the linear switching line in the optimal policy; see Figure 4.

5. *Additional properties of optimal trajectories with linear costs.* Recall that in the case of linear holding costs the switching surfaces are halfspaces. Cosnider two starting states q and aq, where a is a positive constant. It is clear that aq lies in the same region in state space (as defined by

the switching halfspaces), and thus

$$v_q^*(0) = v_{aq}^*(0).$$

In fact pursuing that observation a little further we get that for $t \geq 0$

$$v_{aq}^*(t) = v_q^*(t/a) \quad \text{and} \quad z_{aq}^*(t) = z_q^*(t/a). \tag{III.30}$$

Hence, we only need to compute the optimal trajectories for initial conditions on a ball, say $|q| = 1$, and then use (III.30) to come with the optimal allocations process and optimal trajectory starting from aq for any $a > 0$.

III.6 Numerical Solution of Fluid Optimal Control Problems

The optimization problem described in (III.16) is a continuous time (infinite dimensional) convex program over a polytopic feasible set. Overwhelmingly the most popular choice for the cost rate function g is the linear one. Its popularity is rooted in simple economic arguments, tractability, as well as pure historical reasons. The resulting problem has been addressed quite extensively in the literature.

Linear objectives –Separated Continuous Linear Programs (SCLP). The problem in (III.16) lies in the class of separated continuous linear programs. The essential features for this class of problems are that

- the state dynamics depend linearly on the control and not on the current state itself (i.e., $\dot{z}(t)$ is a linear function of $v(t)$, independent of $z(t)$);
- the state and the control are non-negative;
- the control satisfies polytopic constraints;
- SCLP does not include state polytopic constraints, but, in fact, they can be incorporated without any additional algorithmic or theoretical complication.

We know that the optimal control is *bang-bang*. That is, the optimal control trajectories are piecewise constant and the optimal state trajectories are piecewise linear. This has a the following natural interpretation. In linear programming, the optimum occurs at an extreme point of the polytope of feasible solutions. Similarly, piecewise linear functions are the extreme points within the set of continuous functions, and thus the optimal trajectory for a SCLP again occurs at an extreme point. Consequently, the optimal allocation will have to be piecewise constant. Moreover, extreme points

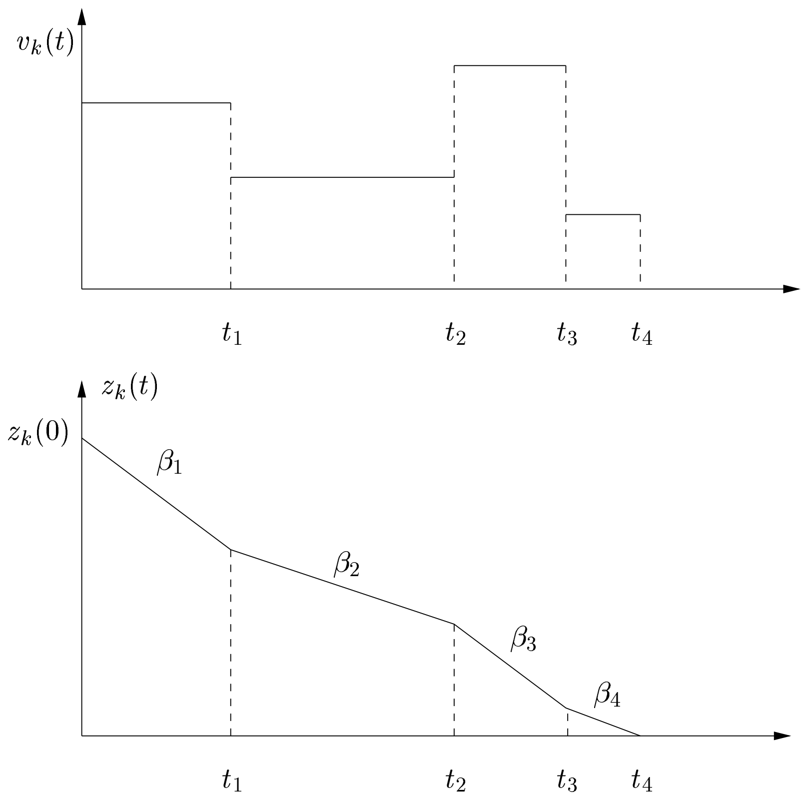

Figure 5: Bang-Bang structure of fluid optimal control solution

will correspond to solutions where the controls will be either at their minimum value, which is 0, or at the maximum value, which is 1; this is the *bang-bang* nature.

Figure 5 depicts an example of control and state trajectories. Breakpoints are at times $t_0, t_1, \dots, t_4$. The controls within each time interval are constant. The corresponding trajectory is piecewise linear; the slopes of each segment are denoted by $\beta_1, \dots, \beta_4$. These slopes are linearly related with the controls. In the sequel we treat them as the decision variables together with the breakpoints $t_0, t_1, \dots, t_4$. Note that instead of the continuum of variables we only have $2K|P|$ now, where $|P|$ is the number of breakpoints. We have,

$$\begin{aligned}\int_0^T h'z(t)dt &= \sum_k h_k \,(\text{area under } z_k \text{ curve}) \\ &= \sum_k h_k \left(\sum [\text{area of trapezoids between } [t_{i-1}, t_i) \text{ of slope } \beta_i] \right).\end{aligned}$$

Note that

$$z(t_i) \text{ is a linear function of } t_0, t_1, \dots, t_i,\ \beta_1, \dots, \beta_i, \text{ and } z(0).$$

As a result the area in each trapezoid will be is a non-convex quadratic function of the decision variables. The SCLP has been reduced into a low dimensional non-convex quadratic program. This reduction has been exploited, and very efficient algorithms have been proposed in the recent literature. This method, however, does not extend to the case of convex (but non-linear) cost rate functions.

Discretization over time. The most natural alternative for the problem in (III.16) is to discretize the system dynamics and solve a finite dimensional convex program.

The discrete dynamics are given by

$$z(r+1) = z(r) + (\alpha - Rv(r))\delta, \quad z(0) = q,$$

and the state and control constraints are

$$z(r) \geq 0, \;\; v(r) \geq 0, Cv(r) \leq \mathbf{1} \text{ for } r = 0, \ldots, \lceil T/\delta \rceil - 1.$$

Extending previous terminology, a pair of sequences (z, v) that satisfy the above set of conditions will be denoted by $(z, v) \in FM_\delta$. The associated discrete time optimal control problem is:

$$\bar{V}_\delta^g(q) = \min_v \left\{ \sum_{r=0}^{\lceil T/\delta \rceil - 1} g(z(r))\delta : (z, v) \in FM_\delta \right\}. \tag{III.31}$$

For the optimal control problem at hand, one can show that for any initial condition q, $\bar{V}_\delta^g(q) \to \bar{V}^g(q)$, as $\delta \to 0$. In practice, suitable choices for δ depend on the arrival rates α and feedback matrix R.

The convexity of $g(\cdot)$, implies that (III.31) is a convex optimization problem over a polytopic constraint set. It can be rewritten in the form

$$\begin{aligned} \bar{V}_\delta^g(q) = \quad & \min \quad t \\ & \text{s.t.} \quad Av \leq b, \quad G^\delta(v) \leq t, \quad v \geq 0 \end{aligned}$$

where $N = \lceil T/\delta \rceil$, $G^\delta(v) = \sum_{r=0}^{N-1} g(z(r))$, and $A,\ b$ describe the polytopic constraint on the control vectors $\{v(0), \ldots, v(N-1)\}$ defined by:

$$A \;=\; \begin{bmatrix} R\delta & & \\ \vdots & \ddots & \\ R\delta & \cdots & R\delta \\ C & & \\ & \ddots & \\ & & C \end{bmatrix}, \; b \;=\; \begin{bmatrix} q + \alpha\delta \\ \vdots \\ q + \alpha N\delta \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

The value of (III.31) depends on the initial condition through the function $G^\delta(v)$, which is the sum of the $g(z(r))$ terms, and through the vector b, which can be expressed in the form $b = Dq + \tilde{b}$ for the obvious choice of D and $\tilde{b}$. The Lagrangian associated with this problem is:

$$L(t, v, \zeta, \nu) = t + \nu^T(Av - b) + \zeta(G^\delta(v) - t) \tag{III.32}$$

By minimizing over t and v, we see that $\zeta = 1$ and that the required constraint for the dual problem to be finite is:

$$A^T\nu \geq -\nabla_v G^\delta. \tag{III.33}$$

¿From the expression for the Lagrangian we can also calculate $\nabla\bar{V}^g(q)$. That is,

$$\frac{\partial L}{\partial q} = \frac{\partial L}{\partial b}\frac{\partial b}{\partial q} + \frac{\partial L}{\partial G^\delta}\frac{\partial G^\delta}{\partial q} = -\nu^T D + \nabla_q G^\delta \tag{III.34}$$

This can be rewritten as

$$\nabla\bar{V}^g_\delta(q) = \delta\left(-\nu^T D + \sum_{r=0}^{N-1} \nabla g(q + \alpha r\delta - \delta R\sum_{j=0}^{r-1} v_j^*)\right)$$

where v^* is the optimal control. Again, additional polytopic constraints on the state and control can be easily included.

The caveat of this method is that the number of variables grows proportionally to the horizon T and inversely proportional to the discretization step δ. This will lead to large optimization problems that are better suited for off-line computation.

Of course, for the case of linear cost rate function we have seen that we only need to compute the optimal solution starting from a small initial condition and then use (III.30) to "zoom out" to the desired degree. Moreover, we have a bound on the control horizon T required to reach the optimal solution that will also get smaller if we only need to compute optimal solutions starting from small initial conditions. Hence, the dimensionality problem is not severe in the case of linear cost rate functions, and, in fact, it can be used very efficiently towards the solution of (III.16).

Value function approximation method. In many applications the "thinking time" allowed in computing the solutions of these fluid optimization problems is limited, and most computation should be carried through off-line. This is important consideration in order to construct solutions that scale gracefully with problem size. In these cases, it is typical to sample the state space at a sequence of points, say $\{q_1, \ldots, q_m\}$, where the solution of the fluid optimal control problem is computed. Let

$$V_i = \bar{V}^g(q_i) \quad \text{and} \quad f_i = \nabla\bar{V}^g(q_i), \quad \text{for } i = 1, \ldots, m.$$

The following functional approximation to the value function $\bar{V}^g$ can be used. First, observe that the objective in (III.16) is convex in the initial condition q and the control sequence $v(\cdot)$. This can be shown using the convexity of g and the observation that given any two feasible controls $v_1(\cdot)$ and $v_2(\cdot)$, the control $v(t) = \nu v_1(t) + (1-\nu)v_2(t)$, for $\nu \in (0,1)$, is also feasible. It is easy to conclude that $\bar{V}^g$ is convex in q and thus

$$\bar{V}^g(q) \geq V_i + f_i'(q - q_i), \quad \text{for } i = 1, \ldots, m.$$

A piecewise linear approximation can be constructed by

$$\hat{V}_g(q) = \max_{1 \leq i \leq m} V_i + f_i'(q - q_i), \tag{III.35}$$

which also serves as a lower bound to the value function of (III.16); that is, $\hat{V}_g(q) \leq \bar{V}^g(q)$ for all $q \geq 0$. The corresponding approximation of $\nabla \bar{V}^g$ is equal to the f_i that corresponds to the maximizer of (III.35). This approximation is valid for convex cost rate functions. Computing the (approximate) optimal allocation is cheap, and such a solution could be used for on-line applications. (Meanwhile, the approximation is continuously refined by computing more points on a grid.)

III.7 Dynamic Control of Fluid Models

Fluid optimal control policies capture optimal transient behavior in the fluid model. These policies are specified using a dynamic index or reward function. Here we consider other dynamic fluid model control policies that are not derived from any explicit optimization of the form in (III.16). The motivation for such an extension is twofold: first, it is often the case that the solution of the fluid optimization problem, is not fully specified, and only an approximate description is available; and second, starting form any initial condition it could be desirable to use control policies derived either by simpler and more direct methodologies, or by incorporating other considerations not included in the formulation of (III.16) such as additional design and operational specifications.

Minimum time control

In many applications it is either hard to have accurate data regarding the cost rate function g, or one does not have access to the solution of the fluid optimal control problem (III.16). In these cases, a natural alternative is to try to optimize transient performance by minimizing the time to drain the initial backlog without any further considerations regarding the cost structure. *Minimum time* control is intuitively appealing and, as we now see, has a surprisingly simple solution in the fluid model.

Assume that $\rho < \mathbf{1}$. Starting from an initial condition q consider the problem of draining the fluid model in minimum time stated below:

$$\bar{V}_I(z) = \min_{v(\cdot)} \left\{ \int_0^\infty I_{\{z(t)\}} dt : z(0) = q \ \text{ and } (z, v) \in FM \right\}, \tag{III.36}$$

where $I_{\{z(t)\}}$ is the indicator function, $I_{\{z(t)\}} = 1$ if $z(t) \neq 0$ else, $I_{\{z(t)\}} = 0$. Given the simple fluid dynamics and using some intuition from the area of minimum time control, one would expect to have bang-bang optimal controls that yield piecewise linear optimal trajectories and a value function $\bar{V}_I$ that is piecewise linear in the initial condition q. We know show that this is true. In fact, the minimum-time control trajectory has already been defined in the proof of Proposition III.4.1, and it corresponds to *linear translation* from q to the origin.

Specifically, let $t^*(q)$ be the minimum draining time under some feasible control. Then,

$$z(t^*(q)) = q + \alpha t^*(q) - R\bar{T}(t^*(q)) = 0$$

which implies that $\bar{T}(t^*(q)) = R^{-1}q + R^{-1}\alpha t^*(q)$. The capacity constraints imply that

$$\begin{aligned} C\bar{T}(t^*(q)) \leq \mathbf{1}t^*(q) \quad &\Rightarrow \quad CR^{-1}q \leq (\mathbf{1} - \rho)t^*(q) \\ &\Rightarrow \quad t^*(q) \geq \max_{1\leq j\leq J} \frac{(CR^{-1}q)_j}{1-\rho_j}. \end{aligned} \tag{III.37}$$

In Proposition III.4.1 it was proved that the control trajectory $\hat{v}(t) = R^{-1}\alpha + R^{-1}q/t^*(q)$ for $t \leq t^*(q)$ and $\hat{v}(t) = R^{-1}\alpha$ for $t > t^*(q)$ achieves the bound in (III.37).

The value function will be $\bar{V}_I(z) = t^*(q)$, which is indeed piecewise linear as it was argued above.

For example, in the Rybko-Stolyar network and starting from $q = [1, 0, 0.5, 1]$, the minimum time control will drain the system in $t^*(q) = 7$ time units. However, the cost accrued under this control is equal to 8.75, which is 21% suboptimal in comparison to the optimal achievable cost in the fluid model; this can be computed from Figure 2 to be 7.22. This result should not be very discouraging, since the derivation of this policy did not involve any computation and could be very practical for large networks.

Reward maximizing policies

Motivated by the structure of (III.29) we now define a family of control policies for fluid models. Specifically, given a dynamic reward rate function $r(\cdot)$, we define a greedy control policy using the following instantaneous resource allocation rule:

$$v(t) \in \operatorname*{argmax}_{v\in\mathcal{V}(q(t))} r(q(t))'v. \tag{III.38}$$

That is, at any point in time the controller allocates resource usage in order to greedily maximize the "instantaneous reward rate" according to (III.38). This policy mimics the fluid optimal control law described in (III.29), with the optimal reward rate function r^g replaced by an arbitrary reward function that has the following interpretation: it associates with each queue length vector q and each job class k a positive value $r_k(q)$, which is treated as a reward rate for time devoted to processing class k jobs. As it was explained earlier, the relative magnitudes of the reward rates induce a dynamic priority rule among classes at the same server.

We replace the optimal reward rate function r^g by an arbitrary reward rate function r that is continuous and further satisfies a non-idling and a polynomial growth condition. The continuity will be needed later on, and the growth and non-idling conditions will be used in order to show stability. In particular we consider the class of reward rate functions $r(\cdot)$ defined on $\mathbf{R}_+^K$, which are real valued, strictly positive, and continuous functions, where each component of which satisfies the growth condition

$$c_1 \le r_k(q) \le c_2 + |q|^{c_3} \quad \text{for some} \;\; c_1,\; c_2,\; c_3 > 0 \;\; \text{and} \; k = 1, \ldots, K. \tag{III.39}$$

We start by establishing that the positivity restriction on the reward rate functions implies the following non-idling property for fluid models under the policy (III.38); the proof is left as a homework problem.

Proposition III.7.1 *The fluid solutions of the set of equations (III.12), (III.13) and (III.38) associated with a strictly positive reward function $r(\cdot)$ are non-idling in the sense that $(Cv(t))_k = 1$ whenever $(Cz(t))_k > 0$.*

Therefore, reward rate functions that satisfy (III.39) can be interpreted as non-idling dynamic priority rules with respect to their fluid behavior. Note, that nothing has been -or can be- said for any queueing network behavior, since for now we have not identified a scheduling policy that achieves (III.29) or (III.38) as their fluid limit. These fluid models are only studied in isolation. Later on, we will take on the problem of defining scheduling policies in the stochastic networks that achieve this fluid limiting behavior.

The specific form of the growth restriction in (III.39) is required for the proof of stability of these models; it could be relaxed at the expense of a more complicated proof. For the appropriate choice of the constant c_3, (III.39) can be made consistent with (III.14). Note, however, that the positivity (or non-idling) restriction need not be satisfied in general by r^g, since optimal policies might allow idling in some cases; recall the 2-station in tandem example. As a result, condition (III.39) does not include all the reward rate functions r^g derived earlier.

Finally, in the case of a constant reward vector, (III.38) simplifies to

$$v(t) \;\in\; \operatorname*{argmax}_{v \in \mathcal{V}(q(t))} \; r'v. \tag{III.40}$$

The constant reward vector will induce a static priority rule according to the relative magnitudes of the various reward rates in the sense described above.

Assuming that $\rho < \mathbf{1}$ we now prove stability of these policies by considering two separate cases when r is a constant vector, and when $r(\cdot)$ is dynamic. Stability is proved by identifying a Lyapunov function with the appropriate negative drift. The case of constant reward rate vectors is addressed in the homework, where you prove the following result:

Proposition III.7.2 *Assuming that $\rho_i < 1$ at every station i, the fluid model described by equations (III.12), (III.13) and (III.40) is stable.*

One last observation in the context of constant reward rate vectors is that (III.40) can be rewritten in the form

$$v(t) \in \underset{v \in \mathcal{V}(q(t))}{\operatorname{argmin}} \frac{dV(q(t))}{dt}. \tag{III.41}$$

Equation (III.41) provides a greedy interpretation of the associated fluid model, in the sense that resource usage is instantaneously allocated in order to minimize a linear objective defined by the linear Lyapunov function $V(q(t))$. The interpretation of the constant vector r as a static priority rule together with equation (III.41) illustrate a relation between linear Lyapunov functions, static priorities, and constant reward rate vectors.

Next, we consider the case of a dynamic reward rate function.

Proposition III.7.3 *Assuming that $\rho_i < 1$ at every station i, the fluid model described by equations (III.12), (III.13) and (III.38) is stable.*

Proof. *Just a sketch.* Let $g(q(t)) = R^{-T}r(q(t))$. Since, R^{-1} is componentwise non-negative it is clear that $g(q(t)) > 0$, for all $q(t) \neq 0$. Furthermore, without loss of generality we can assume that the following normalization condition is true: $1 \leq g(q) \leq b + |q|^\gamma$ for some constant $b > 0$. Define the functional

$$V(q(t)) = -\int_t^\infty e^{-\zeta(\tau - t)} g(q(\tau))' \dot{q}(\tau) d\tau, \tag{III.42}$$

where $\zeta > 0$ is a discount factor. The functional $V(\cdot)$ can be interpreted as an exponentially weighted energy function for the fluid model. The drift of this functional is given by

$$\frac{dV(q(t))}{dt} = g(q(t))'\dot{q}(t). \tag{III.43}$$

The remainder of the proof establishes that this function has the desired negative drift. (The polynomial growth of $r(\cdot)$ is used to prove finiteness of $V(\cdot)$ under discounting.) $\diamond$

An example of a reward rate function. Consider optimal network control problems under linear cost criteria. In these cases, a sensible starting point, in choosing a control policy for these networks, is to try to enforce the priority ranking that at any point in time strives to maximize cost draining out of the system; this is a generalization of the celebrated "$c\mu$" rule. Assuming that the linear holding costs are denoted by h_k for each job class k, this greedy policy strives to minimize $h'(\alpha - Rv)$ of the admissible controls $v \in \mathcal{V}(q(t))$. The corresponding policy is precisely (III.40) for the choice $r = R'h$. Finally, in the case of non-linear costs described by the cost rate function $g(q(t))$, the corresponding dynamic reward rate function will be defined by $r(q(t)) = R'\nabla g(q(t))$.

In the case of the Rybko-Stolyar example analyzed earlier, this reward rate vector will be $r = R'\mathbf{1} = [0, 1.5, 0, 1.5]'$. This, indeed, defines the LBFS static priority ranking that was shown to be optimal in the fluid model in conjunction with some boundary modifications that prevent idleness while maximizing cost draining out of the system when some of the classes have been depleted. It is interesting to note that the fluid model associated with the static priority policy of LBFS can actually be unstable. What is the difference between the static LBFS and its implementation in (III.40)?

III.8 Notes and References

A detailed discussion of the derivation of fluid models and their properties can be found in the notes by Dai [Dai98]; see also Dai [Dai95], Dai and Weiss [DW96], and Bramson [Bra98] for additional comments.

The complexity analysis in Papadimitriou and Tsitsiklis [PT96], considered a similar class of scheduling problems in multiclass networks to the ones studied here; the equivalent for our class of problems easily follows. A detailed discussion of complexity issues in performance and control of queueing network can be found in Bertsimas [Ber95b, section 7].

The fluid model approach to stochastic network control problems has been studied several researchers in the past. Examples can be found in Chen and Yao [CY93], Atkins and Chen [AC95], Avram, Bertsimas and Ricard [ABR95], and Eng, Humphrey and Meyn [EHM96], where heuristic translation mechanisms based on fluid model optimization are described; in fact, related work can be traced back to Newell [New71]. More recent work can be found in Meyn [Mey97] Chen and Meyn [CM98], and Bertsimas and Sethuraman [BS97]. We will return to this framework in the next lecture note.

The Rybko-Stolyar network was studied independently by Kumar and Seidman [KS90] and Rybko and Stolyar [RS92]. The analysis presented here is from Maglaras [Mag99a]. The discussion of fluid optimal control problems in section III.3 is based again on [Mag99a]. Classical results that establish existence of solutions etc. can be found in Luenberger [Lue69]. A detailed -and very good- exposition of dynamic programming can be found in Bertsekas [Ber95a, section 3.2]. Section III.4 is from [Mag99a], where stability and the bound on the time to drain the fluid model under the optimal control were established. Most of the comments in section III.5 are from Pullan [Pul93, Pul95] and Avram, Bertsimas and Ricard [ABR95]. The two station tandem network example is from [ABR95], where the optimal policy was computed by writing the dynamic programming equations and then directly applying the maximum principle. The observation in III.5.5 is new, but is certainly known to other researchers as well.

The discussion on SCLPs is based on Luo and Bertsimas [LB96]. The basic references in this area are the papers due to Pullan [Pul93, Pul95, Pul96], Luo and Bertsimas [LB96], and the work of Weiss [Wei95, Wei97] that addresses specialized examples of SCLPs that have simpler solutions. Despite the success of these algorithms in the case of linear holding costs, there is still need for them to be extended to more general cost structures. For example, as it was argued by Van Mieghem [VM95], convex increasing delay costs provide a more accurate representation of "congestion costs." In this case the resulting problems would be non-linear, yet convex, and the most natural solution technique relies on discretization over time and explicit computation of the resulting finite dimensional convex program. A wide range of such problems can be addressed with very efficient interior point algorithms that perform very well both in theory and in practice. An extensive list of such problems can be found in the papers by Vandenberghe, Boyd and Wu [VBW98], Lobo, Vandenberghe, Boyd and Lebret [LVBL97], and Alizadeh and Schmieta [AS97]. Similar optimal control problems, especially with quadratic costs, have been studied extensively in the context of Model Predictive Control [Cla94], and they should provide a good starting reference for the fluid optimization problems of interest here.

The derivation of the minimum time control is from [Mag99a]. It is a simple generalization of the result due to Weiss [Wei96] for the case of re-entrant line (all arrivals go to class 1, and work flows in a deterministic route from class 1 to class 2 etc – work can visit the same station many times). The description, interpretation and basic results about reward maximizing policies are from Maglaras [Mag99b]. Roughly speaking, greedily optimizing over any reward rate function that satisfies a non-idling constraint (with respect to their fluid model behavior) gives a stable fluid control policy. This whole family of policies, parametrized by the reward rate function $r(\cdot)$ will be used later on to design stable scheduling policies for stochastic networks.

References

[ABR95] 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.

[AC95] D. Atkins and H. Chen. Performance evaluation of scheduling control of queueing networks: fluid model heuristics. *Queueing Systems*, 21:391–413, 1995.

[AS97] F. Alizadeh and S. H. Schmieta. Optimization with semidefinite, quadratic and linear constraints. Technical report, RUTCOR, Rutgers University, November 1997.

[Ber95a] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Vol.1.* Athena Scientific, Belmont, Massachusetts, 1995.

[Ber95b] D. Bertsimas. The achievable region method in the optimal control of queuing systems; formulations, bounds and policies. *Queueing Systems*, 21:337–389, 1995.

[Bra98] M. Bramson. Stability of two families of queueing networks and a discussion of fluid limits. *Queueing Systems*, 28:7–31, 1998.

[BS97] D. Bertsimas and J. Sethuraman. Optimal control of queueing networks: A singular pertubation approach. In *9th INFORMS Applied Probability Conference*, July 1997.

[Cla94] D. Clarke. *Advances in Model Predictive Control.* Oxford University Press, 1994.

[CM98] R-R. Chen and S. P. Meyn. Value iteration and optimization of multiclass queueing networks. *Queueing Systems*, 1998.

[CY93] H. Chen and D. Yao. Dynamic scheduling of a multiclass fluid network. *Oper. Res.*, 41(6):1104–1115, December 1993.

[Dai95] 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.

[Dai98] J. Dai. Stability of fluid and stochastic processing networks, 1998. Lecture notes.

[DW96] J. G. Dai and G. Weiss. Stability and instability of fluid models for certain re-entrant lines. *Math. Oper. Res.*, 21:115–134, 1996.

[EHM96] D. Eng, J. Humphrey, and S. P. Meyn. Fluid network models: Linear programs for control and performance bounds. In *13th World Congress of International Federation of Automatic Control, San Francisco*, 1996.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and Convergence.* Wiley, New York, 1986.

[KS90] 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, March 1990.

[LB96] X. Luo and D. Bertsimas. A new algorithm for state constrained separated linear programs. Technical report, MIT, 1996.

[Lue69] D. G. Luenberger. *Optimization By Vector Space Methods.* John Wiley & Sons, New York, N. Y., 1969.

[LVBL97] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 1997. To appear.

[Mag99a] C. Maglaras. Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality. *Ann. Appl. Prob.*, 1999. Submitted.

[Mag99b] C. Maglaras. Dynamic scheduling in multiclass queueing networks: Stability under discrete-review policies. *Queueing Systems*, 1999. Forthcoming.

[Mey97] 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.

[New71] G. F. Newell. *Applications of Queueing Theory.* Chapman and Hall, 1971.

[PT96] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of optimal queueing network control. *Math. Oper. Res.*, 1996.

[Pul93] M. C. Pullan. An algorithm for a class of continuous linear programs. *SIAM J. on Control and Optimization*, 31(6):1558–1577, November 1993.

[Pul95] M. C. Pullan. Form of optimal solutions for separated continuous linear programs. *SIAM J. on Control and Optimization*, 33(6):1952–1977, November 1995.

[Pul96] M. C. Pullan. A duality theory for separated continuous linear programs. *SIAM J. on Control and Optimization*, 34(3):931–965, May 1996.

[RS92] 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.

[Sto95] A. L. Stolyar. On the stability of multiclass queueing networks: A relaxed sufficient condition via limiting fluid processes. *Markov Processes and Related Fields*, 1:491–512, 1995.

[VBW98] L. Vandenberghe, S. Boyd, and S.-P. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM J. on Matrix Analysis and Applications*, April 1998. To appear.

[VM95] J. A. Van Mieghem. Dynamic scheduling with convex delay costs: The generalized $c\mu$ rule. *Ann. Appl. Prob.*, 5:809–833, 1995.

[Wei95] G. Weiss. On optimal draining of re-entrant fluid lines. In F. Kelly and R. Williams, editors, *Stochastic Networks*, volume 71, pages 91–103. Proceedings of the IMA, 1995.

[Wei96] G. Weiss. Optimal draining of fluid re-entrant lines: some solved examples. In F. Kelly, S. Zachary, and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, pages 19–34. Oxford University Press, 1996.

[Wei97] G. Weiss. Algorithm for minimum wait draining of two station fluid re-entrant line. Technical report, Department of Statistics, The University of Haifa, 1997.