Stability and Asymptotic Optimality of $h$-MaxWeight Policies

Sean Meyn
Department of Electrical and Computer Engineering
University of Illinois & the Coordinated Science Laboratory

NSF support: ECS 05-23620 and DARPA ITMANET
Outline

I Models & Background

II $h$-MaxWeight Policies

III Heavy Traffic

Conclusions
I Modeling & Control 34

4 Scheduling 99
  4.1 Controlled random-walk model ............................. 101
  4.2 Fluid model ................................ 109
  4.3 Control techniques for the fluid model .................. 116

III Stability & Performance 318

9 Optimization 374
  9.4 Optimality equations .................................. 392
  9.6 Optimization in networks ................................. 408

Models & Background
Controlled Random-Walk Model

\[ Q(k+1) = Q(k) + B(k+1)U(k) + A(k+1), \quad Q(0) = x \]

Statistics & topology:

\[
B(k) = \begin{bmatrix}
-S_1(k) & 0 & 0 & 0 \\
S_1(k) & -S_2(k) & 0 & 0 \\
0 & 0 & -S_3(k) & 0 \\
0 & 0 & S_3(k) & -S_4(k)
\end{bmatrix}
\]

\[
A(k) = \begin{bmatrix}
A_1(k) \\
0 \\
A_3(k) \\
0
\end{bmatrix}
\]

Constituency constraints:

\[
CU(k) \leq 1 \\
U(k) \geq 0 \\
C = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}
\]

- Lippman 1975
- Henderson & M. 1997
Fluid Model & Workload

Fluid model captures mean-flow:

\[ q(t) = x + Bz(t) + \alpha t, \quad t \geq 0 \]
\[ q(0) = x \]

\[ B = E[B(k)] = \begin{bmatrix} -\mu_1 & 0 & 0 & 0 \\ \mu_1 & -\mu_2 & 0 & 0 \\ 0 & 0 & -\mu_3 & 0 \\ 0 & 0 & \mu_3 & -\mu_4 \end{bmatrix} \]

\[ \alpha = E[A(k)] = \begin{bmatrix} \alpha_1 \\ 0 \\ \alpha_3 \\ 0 \end{bmatrix} \]

Workload and load parameters:

\[ \xi^1 = \begin{bmatrix} m_1 \\ 0 \\ m_4 \\ m_4 \end{bmatrix}, \quad \xi^2 = \begin{bmatrix} m_2 \\ m_2 \\ m_3 \\ 0 \end{bmatrix} \]
\[ \rho_1 = m_1\alpha_1 + m_4\alpha_3 \]
\[ \rho_2 = m_2\alpha_1 + m_3\alpha_3 \]

with \( m_i = \mu_i^{-1} \)

- Newell 1982, Vandergraft 1983
- Perkins & Kumar 1989
- Chen & Mandelbaum 1991, Cruz 1991
Myopic Policy: Fluid Model

\[ q(t) = x + Bz(t) + \alpha t \]

Constraints:

\[ \frac{d^+}{dt} q(t) = B\zeta(t) + \alpha \]

\( X \) subset of \( \mathbb{R}_+^\ell \)

\( U(x) \) feasible values of \( \zeta(t) \)

when \( x = q(t) \in X \)

Given: Convex monotone cost function,

\[ c: \mathbb{R}_+^\ell \rightarrow \mathbb{R}_+ \]
Myopic Policy: Fluid Model

\[ \frac{d^+}{dt} q(t) = B\zeta(t) + \alpha \]

Constraints: \( x \) subset of \( \mathbb{R}_+^\ell \)

\[ U(x) \text{ feasible values of } \zeta(t) \]
when \( x = q(t) \in X \)

Given: Convex monotone cost function,

\[ c: \mathbb{R}_+^\ell \rightarrow \mathbb{R}_+ \]

\[ \arg\min_{u \in U(x)} \frac{d^+}{dt} c(q(t)) = \arg\min_{u \in U(x)} \langle \nabla c(x), Bu + \alpha \rangle \]
Myopic Policy: CRW Model

\[ Q(k+1) - Q(k) = B(k+1)U(k) + A(k+1) \]

Constraints: \( X_\diamond \) subset of \( \mathbb{R}_+^\ell \) (lattice constraints, etc.)
\[ U_\diamond(x) \] feasible values of \( U(k) \)
when \( x = Q(k) \in X_\diamond \)

Given: Convex monotone cost function,
\[ c: \mathbb{R}_+^\ell \rightarrow \mathbb{R}_+ \]
Myopic Policy: CRW Model

\[ Q(k+1) - Q(k) = B(k+1)U(k) + A(k+1) \]

Constraints: \( X_\diamond \) subset of \( \mathbb{R}^\ell_+ \) (lattice constraints, etc.)

\[ U_\diamond(x) \text{ feasible values of } U(k) \]

when \( x = Q(k) \in X_\diamond \)

Given: Convex monotone cost function,

\[ c: \mathbb{R}^\ell_+ \rightarrow \mathbb{R}_+ \]

Myopic policy:

\[ \arg \min_{u \in U_\diamond(x)} E[c(Q(k + 1)) \mid Q(k) = x, U(k) = u] \]
Fluid Model & Myopia

\[ q(t) = x + Bz(t) + \alpha t, \quad t \geq 0 \]

\[ q(0) = x \]

Given: Convex monotone cost function,

\[ c: \mathbb{R}_+^\ell \to \mathbb{R}_+ \]

Myopic policy for fluid model is stabilizing:

\[ q(t) = 0 \quad t \geq T_0 \]
Myopia & Instability

Myopic policy may or may not be stabilizing

Example: Two station model above with linear cost,

\[ c(x) = x_1 + x_2 + x_3 + x_4 \]

Myopic policy for CRW model: Priority to exit buffers
Myopia & Instability

Myopic policy may or may not be stabilizing

Example: Two station model above with linear cost,

\[ c(x) = x_1 + x_2 + x_3 + x_4 \]

Myopic policy for CRW model: Priority to exit buffers

Periodic starvation creates instability
Myopia & Instability

Quadratic Cost

Myopic policy stabilizing for *diagonal* quadratic

Example: Two station model above with,

\[ c(x) = \frac{1}{2}[x_1^2 + x_2^2 + x_3^2 + x_4^2] \]

Myopic policy: Approximated by linear switching curves
Myopia & Instability

Quadratic Cost

Myopic policy stabilizing for \textit{diagonal} quadratic

Example: Two station model above with,

\[ c(x) = \frac{1}{2} [x_1^2 + x_2^2 + x_3^2 + x_4^2] \]

Myopic policy: Approximated by linear switching curves

Condition (V3) holds with Lyapunov function \( V = c \)

For positive constants \( \varepsilon \) and \( \eta \)

\[
PV(x) := E[V(Q(k + 1))|Q(k) = x] \leq V(x) - \varepsilon \|x\| + \eta
\]
Tassiulas considers myopic policy for fluid model

\[
\arg \min_{u \in U(x)} \langle \nabla c(x), Bu + \alpha \rangle
\]
subject to lattice constraints

where \( c(x) = \frac{1}{2} x^T D x \), \( D = \text{diag}(d_1, \ldots, d_\ell) \)
Tassiulas considers myopic policy *for fluid model*

\[
\arg\min_{u \in U_\triangle(x)} \langle \nabla c(x), Bu + \alpha \rangle
\]

subject to lattice constraints

Obtains negative drift: For non-zero \( x \),

\[
\langle \nabla c(x), Bu + \alpha \rangle \leq -\varepsilon \|x\|
\]

Implies (V3) for MaxWeight policy
MaxWeight Policy

Tassiulas considers myopic policy for fluid model

\[
\arg\min_{u \in \mathcal{U}(x)} \langle \nabla c(x), Bu + \alpha \rangle \quad \text{subject to lattice constraints}
\]

Obtains negative drift: For non-zero \( x \),

\[
\langle \nabla c(x), Bu + \alpha \rangle \leq -\varepsilon \|x\|
\]

Implies (V3) for MaxWeight policy

Implies (V3) for myopic policy

since myopic has minimum drift
Control Techniques for Complex Networks

I  Modeling & Control

4.8  MaxWeight and MinDrift ............................................. 145
4.9  Perturbed value function ............................................. 148
4.10 Notes ................................................................. 152
4.11 Exercises .............................................................. 154

II  h-MaxWeight Policies

III  Stability & Performance

8  Foster-Lyapunov Techniques

8.1  Lyapunov functions ................................................... 324
8.4  MaxWeight ............................................................. 342
8.5  MaxWeight and the average-cost optimality equation ........... 348
Why Does MW Work?

Geometric explanation

Define drift vector field (for given policy)

\[ \Delta(x) = \mathbb{E}[Q(k+1) - Q(k) \mid Q(k) = x] = Bu + \alpha \]

MaxWeight policy:

\[ \arg \min_{u \in U(x)} \langle \nabla c(x), \Delta(x) \rangle \]

with \( c \) diagonal quadratic
Why Does MW Work?

\[ \Delta(x) = \text{E}[Q(k + 1) - Q(k) \mid Q(k) = x] \]

Example: Queues in tandem

\[ \Delta(x) = \left( \frac{\alpha_1}{-\mu_2} \right) \]

\[ \Delta(x) = \left( \frac{-\mu_1 + \alpha_1}{-\mu_2 + \mu_1} \right) \]

\[ \Delta(x) = \left( \frac{-\mu_1 + \alpha_1}{\mu_1} \right) \]

MaxWeight policy: serve buffer 1
Why Does MW Work?

\[ \Delta(x) = \mathbb{E}[Q(k + 1) - Q(k) \mid Q(k) = x] \]

Example: Queues in tandem

Key observation: Boundaries of the state space are repelling
Why Does MW Work?

\[ \Delta(x) = \mathbb{E}[Q(k+1) - Q(k) \mid Q(k) = x] \]

Example: Queues in tandem

Key observation: Boundaries of the state space are repelling

Consequence of vanishing partial derivatives on boundary
$h$-MaxWeight Policy

Given: Convex monotone function $h$

Boundary conditions

$$\frac{\partial}{\partial x_j} h(x) = 0 \quad \text{when } x_j = 0.$$
\( h \)-MaxWeight Policy

**Given:** Convex monotone function \( h \)

**Boundary conditions**

\[
\frac{\partial}{\partial x_j} h(x) = 0 \quad \text{when} \quad x_j = 0.
\]

**Economic interpretation:**

*Marginal disutility vanishes for vanishingly small inventory*
\( h\)-MaxWeight Policy

Given: Convex monotone function \( h \)

Boundary conditions

\[
\frac{\partial}{\partial x_j} h(x) = 0 \quad \text{when} \quad x_j = 0.
\]

Economic interpretation:

*Marginal disutility vanishes for vanishingly small inventory*

Condition rarely holds, but we can fix that ...
**$h$-MaxWeight Policy**

Given: Convex monotone function $h_0$ (perhaps violating $\partial$ condition)

Introduce perturbation: For fixed $\theta \geq 1$ and any $x \in \mathbb{R}^\ell_+$

\[
\tilde{x}_i := x_i + \theta(e^{-x_i/\theta} - 1), \quad \text{and} \quad \tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_\ell)^T \in \mathbb{R}^\ell_+
\]
**h-MaxWeight Policy**

Given: Convex monotone function $h_0$ (perhaps violating $\partial \theta$ condition)

Introduce perturbation: For fixed $\theta \geq 1$ and any $x \in \mathbb{R}_+^\ell$

$$\tilde{x}_i := x_i + \theta(e^{-x_i/\theta} - 1), \quad \text{and} \quad \tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_\ell)^T \in \mathbb{R}_+^\ell$$

Perturbed function:

$$h(x) = h_0(\tilde{x}), \quad x \in \mathbb{R}_+^\ell$$

Convex, monotone, and boundary conditions are satisfied
$h$-MaxWeight Policy

Perturbed linear function

$h_0$ linear: *never* satisfies $\partial$ condition

$h$-myopic and $h$-MaxWeight polices stabilizing provided $\theta \geq 1$ is sufficiently large
**h-**MaxWeight Policy

*Perturbed linear function*

\( h_0 \) linear: *never* satisfies \( \partial \) condition

\( h \)-myopic and \( h \)-MaxWeight policies stabilizing provided \( \theta \geq 1 \) is sufficiently large

**Example: Tandem queues**

\[ \Delta(x) = \left( \begin{array}{c} \alpha_1 \\ -\mu_2 \end{array} \right) \]

\[ \bar{q}_2 = -\theta \log \left( 1 - \frac{c_1}{c_2} \right) \]

\[ \Delta(x) = \left( \begin{array}{c} -\mu_1 + \alpha_1 \\ -\mu_2 + \mu_1 \end{array} \right) \]
$h$-MaxWeight Policy

Perturbed value function

$h_0$ minimal fluid value function,

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

$h$-myopic and $h$-MaxWeight polices stabilizing

provided $\theta \geq 1$ is sufficiently large
\( h - \text{MaxWeight Policy} \)

\textit{Perturbed value function}

\( h_0 \) minimal fluid value function,
\[
J(x) = \inf \int_0^\infty c(q(t; x)) \, dt
\]

\( h \)-myopic and \( h \)-MaxWeight polices stabilizing provided \( \theta \geq 1 \) is sufficiently large

Resulting policy very similar to average-cost optimal policy:
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Workload</td>
<td>158</td>
</tr>
<tr>
<td>5</td>
<td>Workload &amp; Scheduling</td>
<td>159</td>
</tr>
<tr>
<td>5.1</td>
<td>Single server queue</td>
<td>160</td>
</tr>
<tr>
<td>5.2</td>
<td>Workload for the CRW scheduling model</td>
<td>163</td>
</tr>
<tr>
<td>5.3</td>
<td>Relaxations for the fluid model</td>
<td>167</td>
</tr>
<tr>
<td>III</td>
<td>Stability &amp; Performance</td>
<td>318</td>
</tr>
<tr>
<td>9</td>
<td>Optimization</td>
<td>374</td>
</tr>
<tr>
<td>10</td>
<td>ODE methods</td>
<td>436</td>
</tr>
<tr>
<td>10.5</td>
<td>Safety stocks and trajectory tracking</td>
<td>462</td>
</tr>
<tr>
<td>10.6</td>
<td>Fluid-scale asymptotic optimality</td>
<td>467</td>
</tr>
</tbody>
</table>
Relaxations & Asymptotic Optimality

Single example for sake of illustration:

Model of Dai & Wang
Relaxations & Asymptotic Optimality

Single example for sake of illustration:

Assume: Homogeneous model
Service rate at Station $i$ is $\mu_i$
Relaxations & Asymptotic Optimality

Homogeneous CRW model:

\[ Q_1(k+1) - Q_1(k) = -S_1(k+1)U_1(k) + A_1(k+1) \]
\[ Q_2(k+1) - Q_2(k) = -S_1(k+1)U_2(k) + S_1(k+1)U_1(k) \]
\[ Q_3(k+1) - Q_3(k) = -S_2(k+1)U_3(k) + S_2(k+1)U_2(k) \]
\[ Q_4(k+1) - Q_4(k) = -S_2(k+1)U_4(k) + S_2(k+1)U_3(k) \]
\[ Q_5(k+1) - Q_5(k) = -S_1(k+1)U_5(k) + S_2(k+1)U_4(k) \]
Relaxations & Asymptotic Optimality

Homogeneous CRW model:

\[ Q_1(k+1) - Q_1(k) = -S_1(k+1)U_1(k) + A_1(k+1) \]
\[ Q_2(k+1) - Q_2(k) = -S_1(k+1)U_2(k) + S_1(k+1)U_1(k) \]
\[ Q_3(k+1) - Q_3(k) = -S_2(k+1)U_3(k) + S_2(k+1)U_2(k) \]
\[ Q_4(k+1) - Q_4(k) = -S_2(k+1)U_4(k) + S_2(k+1)U_3(k) \]
\[ Q_5(k+1) - Q_5(k) = -S_1(k+1)U_5(k) + S_2(k+1)U_4(k) \]

Constituency constraints: \( U_i(k) \in \{0, 1\} \)

\[ U_1(k) + U_2(k) + U_5(k) \leq 1 \quad U_3(k) + U_4(k) \leq 1 \]
Relaxations & Asymptotic Optimality

Workload (units of inventory)

\[ Y_1(k) = 3Q_1(k) + 2Q_2(k) + Q_3(k) + Q_4(k) + Q_5(k) \]

\[ Y_2(k) = 2(Q_1(k) + Q_2(k) + Q_3(k)) + Q_4(k) \]
Relaxations & Asymptotic Optimality

Workload (units of inventory)

\[ Y_1(k) = 3Q_1(k) + 2Q_2(k) + Q_3(k) + Q_4(k) + Q_5(k) \]
\[ Y_2(k) = 2(Q_1(k) + Q_2(k) + Q_3(k)) + Q_4(k) \]

Idleness processes:

\[ \nu_1(k) = 1 - (U_1(k) + U_2(k) + U_5(k)) \]
\[ \nu_2(k) = 1 - (U_3(k) + U_4(k)) \]
Relaxations & Asymptotic Optimality

Workload (units of inventory)

\[ Y_1(k) = 3Q_1(k) + 2Q_2(k) + Q_3(k) + Q_4(k) + Q_5(k) \]
\[ Y_2(k) = 2(Q_1(k) + Q_2(k) + Q_3(k)) + Q_4(k) \]

Idleness processes:

\[ \nu_1(k) = 1 - (U_1(k) + U_2(k) + U_5(k)) \]
\[ \nu_2(k) = 1 - (U_3(k) + U_4(k)) \]

Dynamics:

\[ Y_1(k + 1) - Y_1(k) = -S_1(k + 1) + 3A_1(k + 1) + S_1(k + 1)\nu_1(k) \]
\[ Y_2(k + 1) - Y_2(k) = -S_2(k + 1) + 2A_1(k + 1) + S_2(k + 1)\nu_2(k) \]
Relaxations & Asymptotic Optimality

Workload Relaxation of N. Laws

\[ Y_1(k + 1) - Y_1(k) = -S_1(k + 1) + 3A_1(k + 1) + S_1(k + 1)\iota_1(k) \]

with constraints on idleness process relaxed,

\[ \iota_1(k) \in \{0, 1, 2, \ldots \} \]
Relaxations & Asymptotic Optimality

Workload Relaxation of N. Laws

\[ Y_1(k + 1) - Y_1(k) = -S_1(k + 1) + 3A_1(k + 1) + S_1(k + 1)\nu_1(k) \]

with constraints on idleness process relaxed,

\[ \nu_1(k) \in \{0, 1, 2, \ldots\} \]

Optimization based on the effective cost,

\[ \bar{c}(y) = \min \quad c(x) \]

s. t. \[ 3x_1 + 2x_2 + x_3 + x_4 + x_5 = y \]

\[ x \in \mathbb{Z}^5_+ \quad (+ \text{ buffer constraints}) \]
Asymptotic Optimality

Optimal policy is non-idling for one-dimensional relaxation

Dynamic programing equation solved via *Pollaczek-Khintchine* formula
Asymptotic Optimality
Heavy traffic assumptions

Load is unity for nominal model
Single bottleneck to define relaxation
Cost is linear, and effective cost has a unique optimizer

Model sequence:

\[ A^{(n)}(k) = \begin{cases} 
A(k) & \text{with probability } 1 - n^{-1} \\
0 & \text{with probability } n^{-1}
\end{cases} \]

Load less than unity for each \( n \)
Asymptotic Optimality

\[ h_0(x) = \hat{h}^*(y) + \frac{b}{2} \left( c(x) - \bar{c}(y) \right)^2 \]

\( h \)-MaxWeight policy asymptotically optimal, with logarithmic regret
Asymptotic Optimality

\[ h_0(x) = \hat{h}^*(y) + \frac{b}{2} \left( c(x) - \bar{c}(y) \right)^2 \]

\( h \)-MaxWeight policy asymptotically optimal, with logarithmic regret

\[ \hat{\eta}^* = O(n) \]  optimal average cost for relaxation

\[ \eta \]  average cost under \( h \)-MW policy
Asymptotic Optimality

\[ h_0(x) = \hat{h}^*(y) + \frac{b}{2} \left( c(x) - \bar{c}(y) \right)^2 \]

\( h \)-MaxWeight policy asymptotically optimal, with logarithmic regret

\[ \hat{\eta}^* = O(n) \quad \text{optimal average cost for relaxation} \]

\( \eta \quad \text{average cost under} \ h \text{-MW policy} \)

\[ \hat{\eta}^* \leq \eta \leq \hat{\eta}^* + O(\log(n)) \]
Conclusions
Conclusions

$h$-MaxWeight policy stabilizing under very general conds.

General approach to policy translation. Resulting policy mirrors optimal policy in examples

Asymptotically optimal, with logarithmic regret for model with single bottleneck
Conclusions

$h$-MaxWeight policy stabilizing under very general conds.

General approach to policy translation. Resulting policy mirrors optimal policy in examples

Asymptotically optimal, with logarithmic regret for model with single bottleneck

Future work

Models with multiple bottlenecks?

On-line learning for policy improvement?

[C. Moallemi, B. Van Roy, S. Kumar 2007]
References


