# Heterogeneous Congestion Control Protocols

Steven Low

with **A. Tang**, J. Wang, D. Wei, Caltech
M. Chiang, Princeton

# Outline

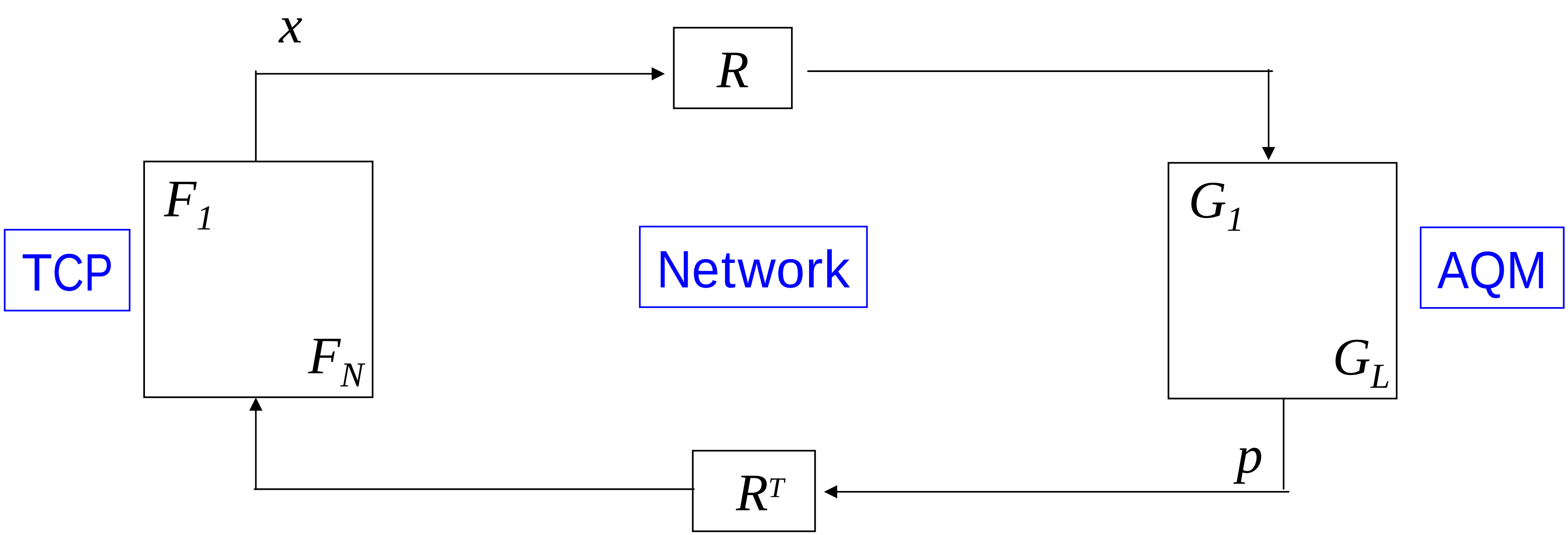
- Review: homogeneous case
- Motivating experiments
- Model
- Equilibrium
  - Existence, uniqueness, local stability
  - Efficiency, fairness
- Slow timescale control

Tang, Wang, Low, Chiang. ToN, 2007
Tang, Wang, Hegde, Low. Telecommunications Systems, Dec 2005
Tang, Wei, Low, Chiang. ICNP, 2006

# Bibliography!!!

- Bibliography!!!

# Network model




$$R_{li} = 1 \quad \text{if source } i \text{ uses link } l$$ ← IP routing

$$x(t+1) = F(R^T p(t),\ x(t))$$ ← Reno, Vegas

$$p(t+1) = G(p(t),\ Rx(t))$$ ← DT, RED, …

# Network model: example

**Reno:**
Jacobson
1989

```
for every RTT          (AI)
{   W += 1   }
for every loss         (MD)
{   W := W/2   }
```

AI

$$x_i(t+1) = \frac{1}{T_i^2} - \frac{x_i^2}{2} \sum_l R_{li}\, p_l(t)$$

MD

$$p_l(t+1) = G_l\left( \sum_i R_{li}\, x_i(t),\ p_l(t) \right)$$

TailDrop

# Network model: example

FAST:

Jin, Wei, Low
2004

```
periodically
{
        W := (baseRTT / RTT) W + α
}
```

$$x_i(t+1) = x_i(t) + \frac{\gamma_i}{T_i}\left(\alpha_i - x_i(t)\sum_l R_{li}\, p_l(t)\right)$$

$$p_l(t+1) = p_l(t) + \frac{1}{c_l}\left(\sum_i R_{li}\, x_i(t) - c_l\right)$$

# Duality model of TCP/AQM

- TCP/AQM
$$x^* = F(R^T p^*,\ x^*)$$
$$p^* = G(p^*,\ Rx^*)$$

- Equilibrium $(x^*, p^*)$ primal-dual optimal:
$$\max_{x \geq 0} \sum U_i(x_i) \quad \text{subject to} \quad Rx \leq c$$
  - $F$ determines utility function $U$
  - $G$ guarantees complementary slackness
  - $p^*$ are Lagrange multipliers

Kelly, Maloo, Tan 1998
Low, Lapsley 1999

Uniqueness of equilibrium
- $x^*$ is unique when $U$ is strictly concave
- $p^*$ is unique when $R$ has full row rank

# Duality model of TCP/AQM

- TCP/AQM $$x^* = F(R^T p^*, x^*)$$ $$p^* = G(p^*, Rx^*)$$
- Equilibrium $(x^*, p^*)$ primal-dual optimal:

$$\max_{x \geq 0} \sum U_i(x_i) \quad \text{subject to} \quad Rx \leq c$$

  - $F$ determines utility function $U$
  - $G$ guarantees complementary slackness
  - $p^*$ are Lagrange multipliers

Kelly, Maloo, Tan 1998
Low, Lapsley 1999

The underlying concave program also leads to simple dynamic behavior

# Duality model of TCP/AQM

- Equilibrium $(x^*, p^*)$ primal-dual optimal:

$$\max_{x \ge 0} \sum U_i(x_i) \quad \text{subject to} \quad Rx \le c$$

Mo & Walrand 2000:

$$U_i(x_i) = \begin{cases} \log x_i & \text{if } \alpha = 1 \\ (1-\alpha)^{-1} x_i^{1-\alpha} & \text{if } \alpha \ne 1 \end{cases}$$

- $\alpha = 1$ : Vegas, FAST, STCP
- $\alpha = 1.2$: HSTCP
- $\alpha = 2$ : Reno
- $\alpha = \infty$ : XCP (single link only)

Low 2003

# Duality model of TCP/AQM

- Equilibrium *(x*,p*)* primal-dual optimal:

$$\max_{x \geq 0} \sum U_i(x_i) \qquad \text{subject to} \quad Rx \leq c$$

Mo & Walrand 2000:

$$U_i(x_i) = \begin{cases} \log x_i & \text{if} \quad \alpha = 1 \\ (1-\alpha)^{-1} x_i^{1-\alpha} & \text{if} \quad \alpha \neq 1 \end{cases}$$

- $\alpha = 0$: maximum throughput
- $\alpha = 1$: proportional fairness
- $\alpha = 2$: min delay fairness
- $\alpha = \infty$: maxmin fairness

Low 2003

# Some implications

- Equilibrium
  - Always exists, unique if $R$ is full rank
  - Bandwidth allocation independent of AQM or arrival
  - Can predict macroscopic behavior of large scale networks
- Counter-intuitive throughput behavior
  - Fair allocation is not always inefficient
  - Increasing link capacities do not always raise aggregate throughput

[Tang, Wang, Low, ToN 2006]

- FAST TCP
  - Design, analysis, experiments

[Jin, Wei, Low, ToN 2007]

# Some implications

- Equilibrium
  - Always exists, unique if $R$ is full rank
  - Bandwidth allocation independent of AQM or arrival
  - Can predict macroscopic behavior of large scale networks
- Counter-intuitive throughput behavior
  - Fair allocation is not always inefficient
  - Increasing link capacities do not always raise aggregate throughput

[Tang, Wang, Low, ToN 2006]

- FAST TCP
  - Design, analysis, experiments

[Jin, Wei, Low, ToN 2007]

# Duality model

- Global stability in absence of feedback delay
  - Lyapunov function
    - Kelly, Maulloo & Tan (1988)
  - Gradient projection
    - Low & Lapsley (1999)
  - Singular perturbations
    - Kunniyur & Srikant (2002)
  - Passivity approach
    - Wen & Arcat (2004)
- Linear stability in presence of feedback delay
  - Nyquist criteria
    - Paganini, Doyle, Low (2001), Vinnicombe (2002), Kunniyur & Srikant (2003)
- Global stability in presence of feedback delay
  - Lyapunov-Krasovskii, SoSTool
    - Papachristodoulou (2005)
  - Global nonlinear invariance theory
    - Ranjan, La & Abed (2004, delay-independent)

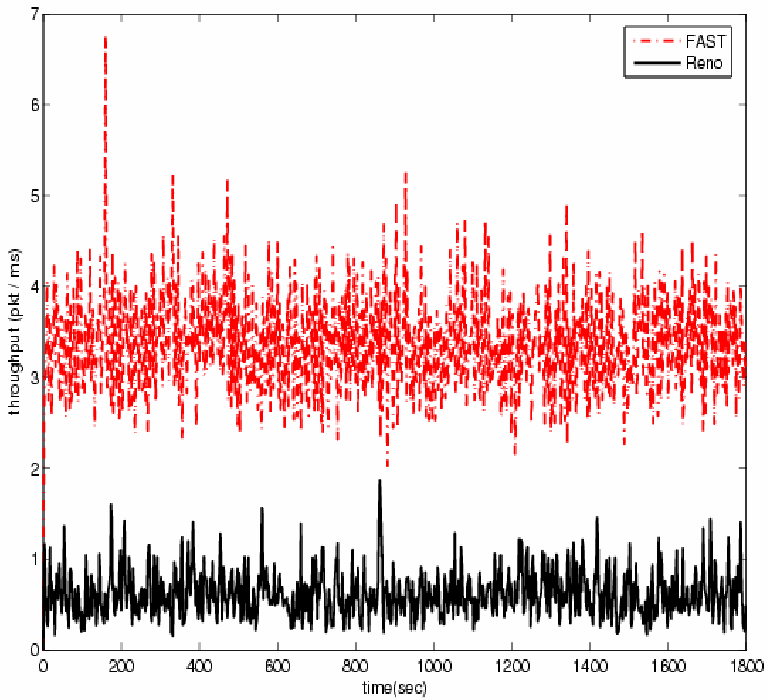
# Outline

- Review: homogeneous case
- Motivating experiments
- Model
- Equilibrium
  - Existence, uniqueness, local stability
  - Efficiency, fairness
- Slow timescale control

# The world is heterogeneous...

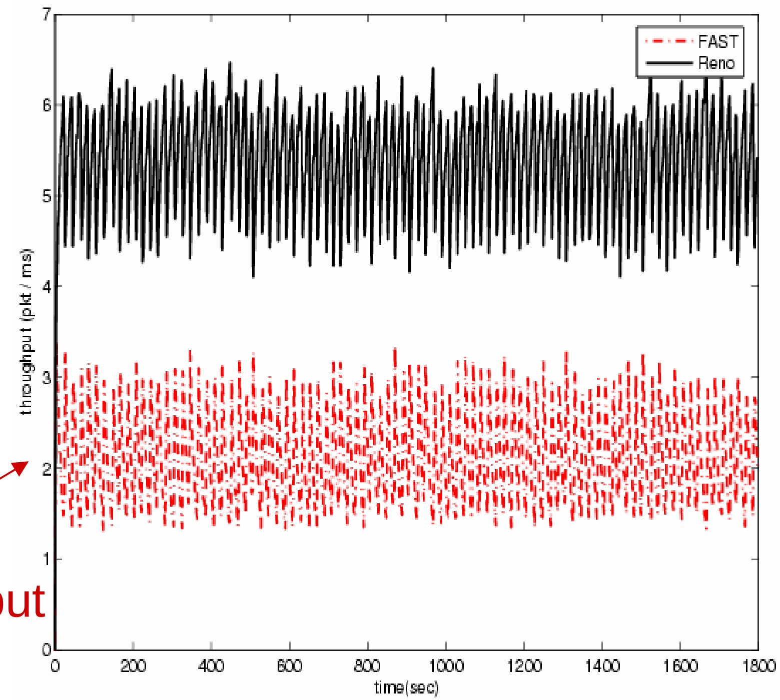
- Linux 2.6.13 allows users to choose congestion control algorithms
- Many protocol proposals
  - Loss-based: Reno and a large number of variants
  - Delay-based: CARD (1989), DUAL (1992), Vegas (1995), FAST (2004), ...
  - ECN: RED (1993), REM (2001), PI (2002), AVQ (2003), ...
  - Explicit feedback: MaxNet (2002), XCP (2002), RCP (2005), ...

# Throughputs depend on AQM



buffer size = 80 pkts

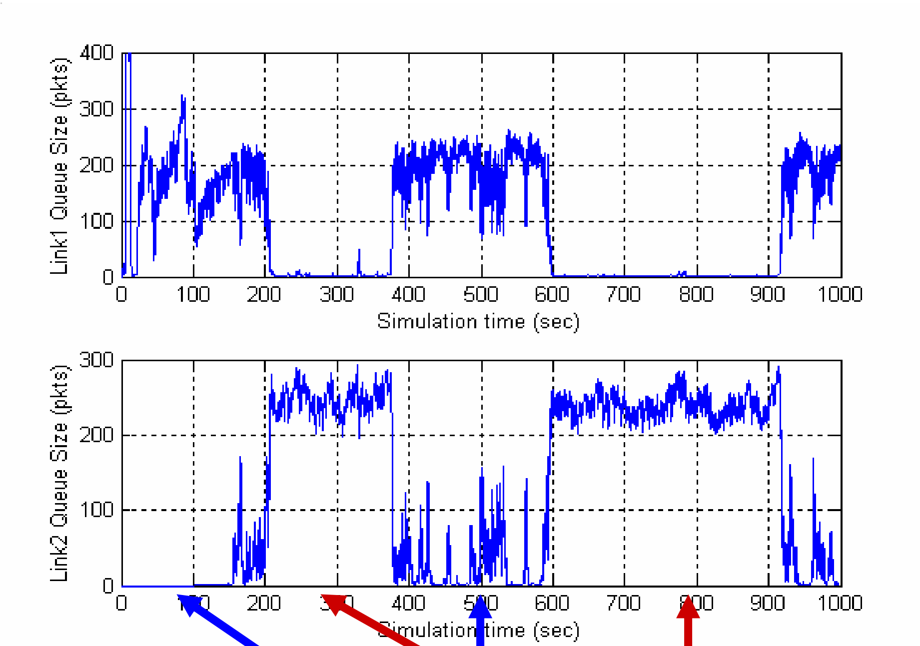
FAST throughput

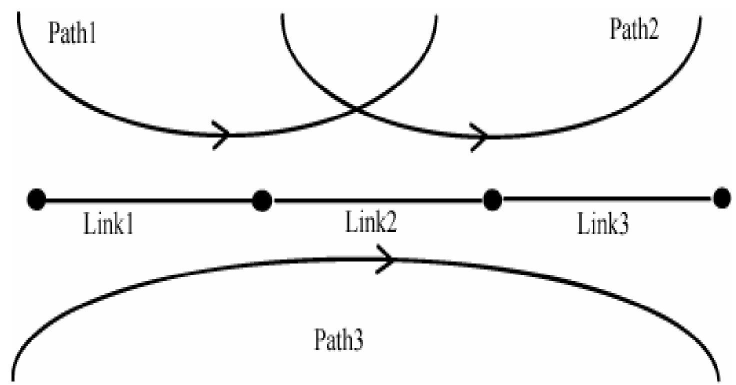


buffer size = 400 pkts

- FAST and Reno share a single bottleneck router
- NS2 simulation
- Router: DropTail with variable buffer size
- With 10% heavy-tailed noise traffic

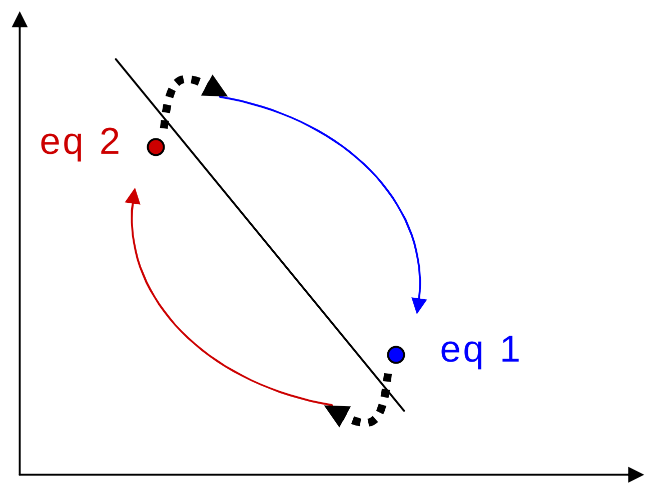
# Multiple equilibria: throughput depends on arrival



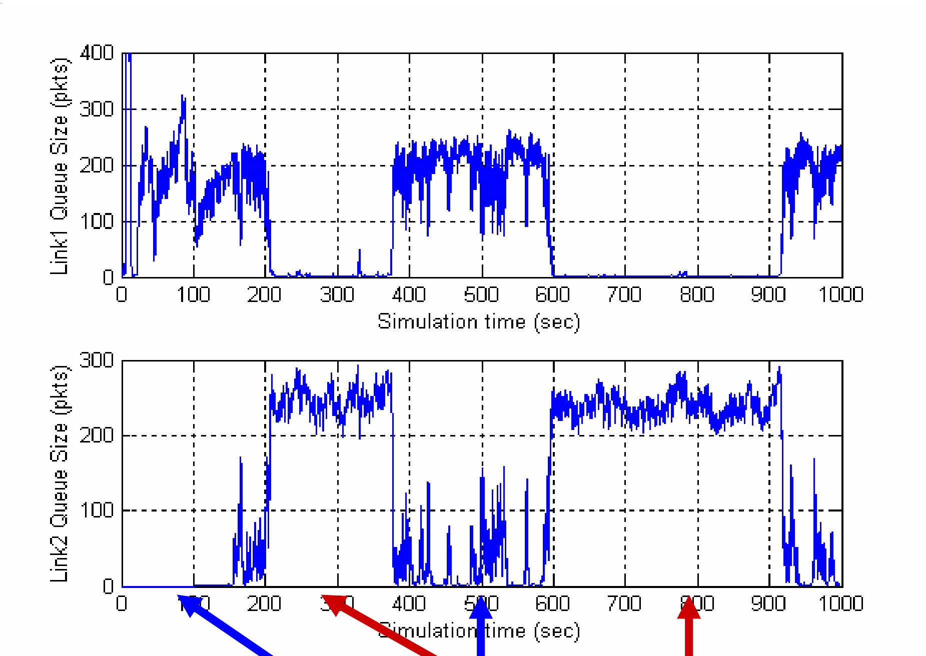

Dummynet experiment

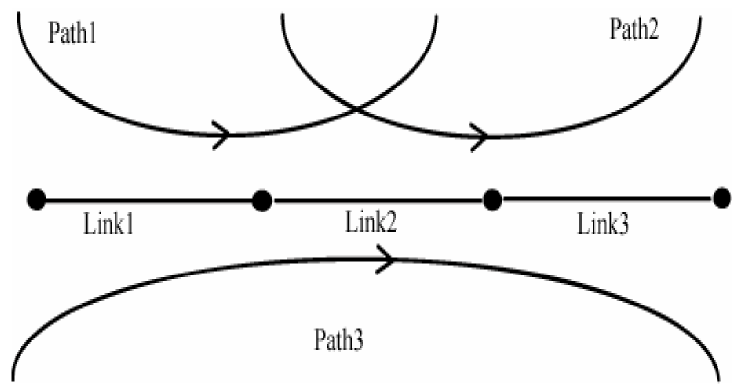
| | eq 1 | eq 2 |
|---|---|---|
| Path 1 | 52M | 13M |
| path 2 | 61M | 13M |
| path 3 | 27M | 93M |




Tang, Wang, Hegde, Low, Telecom Systems, 2005

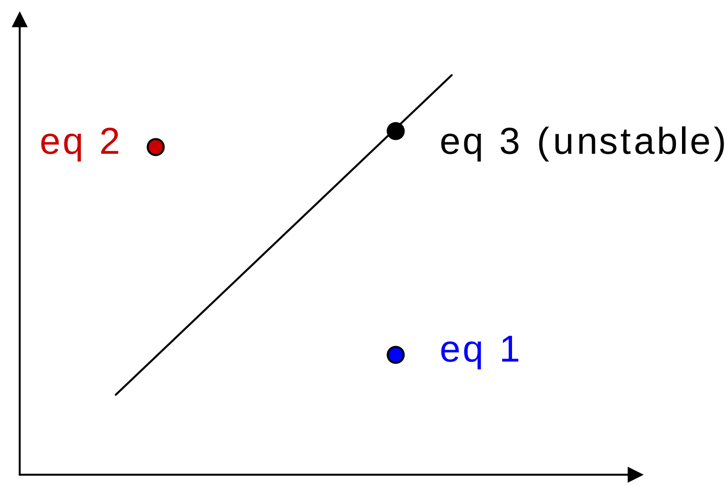
# Multiple equilibria: throughput depends on arrival



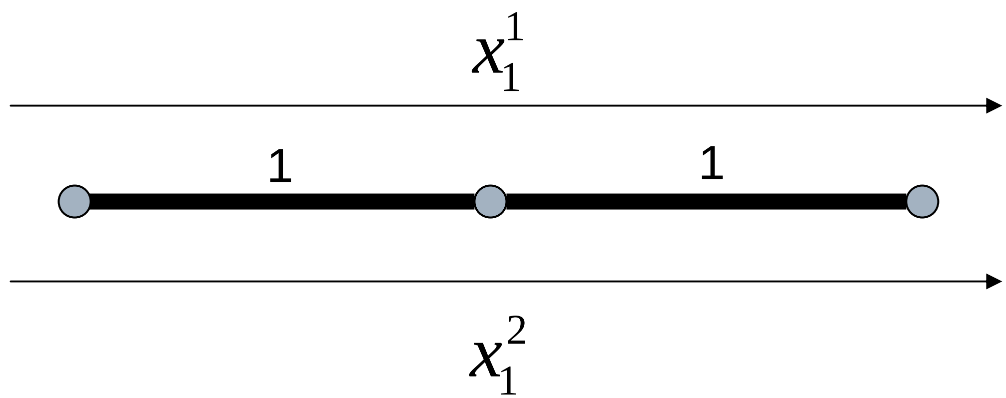

Dummynet experiment

| | eq 1 | eq 2 |
|---|---|---|
| Path 1 | 52M | 13M |
| path 2 | 61M | 13M |
| path 3 | 27M | 93M |




Tang, Wang, Hegde, Low, Telecom Systems, 2005

# Multiple equilibria: single constraint sets




- Smallest example for multiple equilibria
- Single constraint set but infinitely many equilibria
- J=1: prices are non-unique but rates are unique
- J>1: prices and rates are both non-unique

# Some implications

| | **homogeneous** | **heterogeneous** |
|---|---|---|
| **equilibrium** | unique | non-unique |
| **bandwidth allocation on AQM** | independent | dependent |
| **bandwidth allocation on arrival** | independent | dependent |

- Duality model:

$$\max_{x\geq 0} \sum U_i(x_i) \quad \text{s.t. } Rx \leq c \qquad\qquad x_i^* = F_i\left(\sum_l R_{li} p_l^*, x_i^*\right)$$

- Why can't use $F_i$'s of FAST and Reno in duality model?

They use different prices!

$$F_i = x_i + \frac{\gamma_i}{T_i}\left(\alpha_i - x_i \sum_l R_{li} p_l\right)$$

delay for FAST

$$F_i = \frac{1}{T_i^2} - \frac{x_i^2}{2} \sum_l R_{li} p_l$$

loss for Reno

- Duality model:

$$\max_{x \geq 0} \sum U_i(x_i) \quad \text{s.t. } Rx \leq c \qquad x_i^* = F_i\left(\sum_l R_{li} p_l^*, x_i^*\right)$$

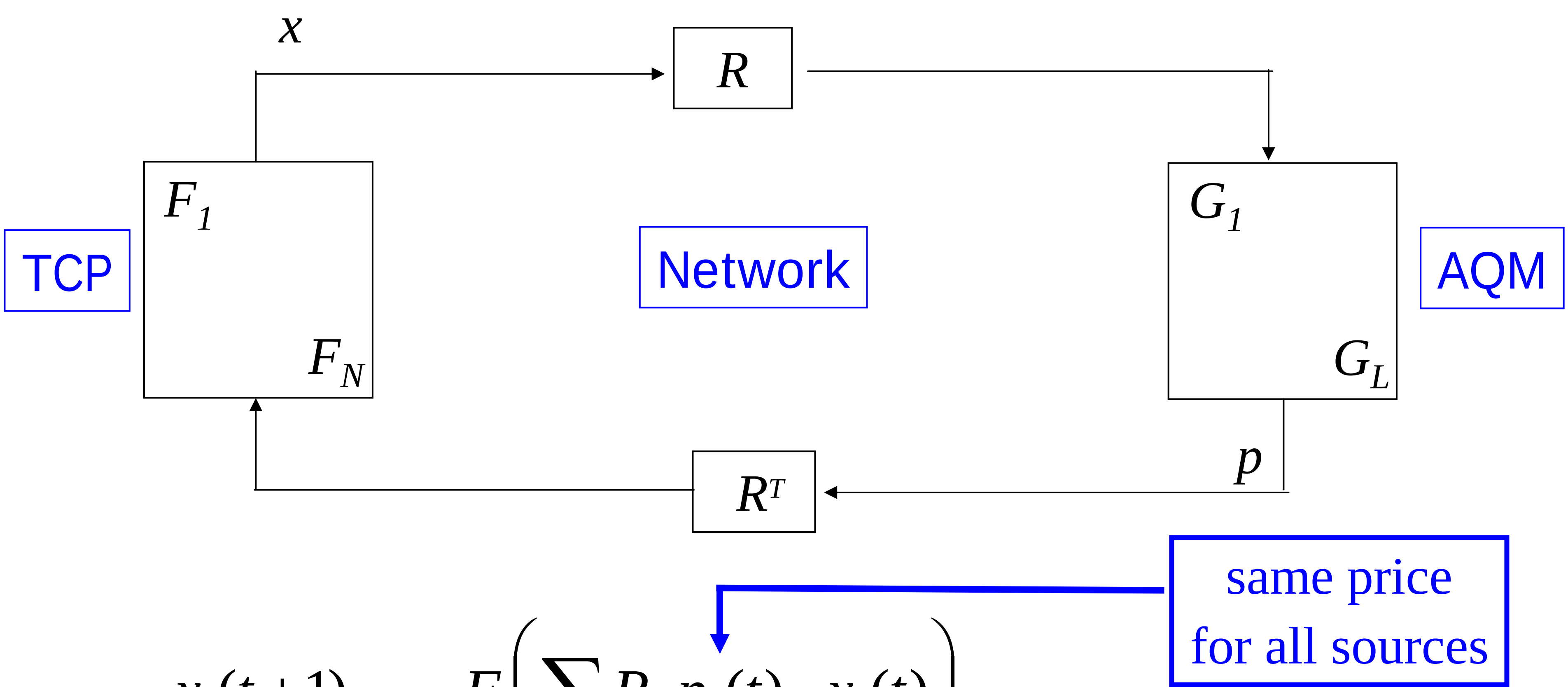
- Why can't use $F_i$'s of FAST and Reno in duality model?

They use different prices!

$$F_i = x_i + \frac{\gamma_i}{T_i}\left(\alpha_i - x_i \sum_l R_{li} p_l\right) \longleftarrow \dot{p}_l = \frac{1}{c_l}\left(\sum_i R_{li} x_i(t) - c_l\right)$$

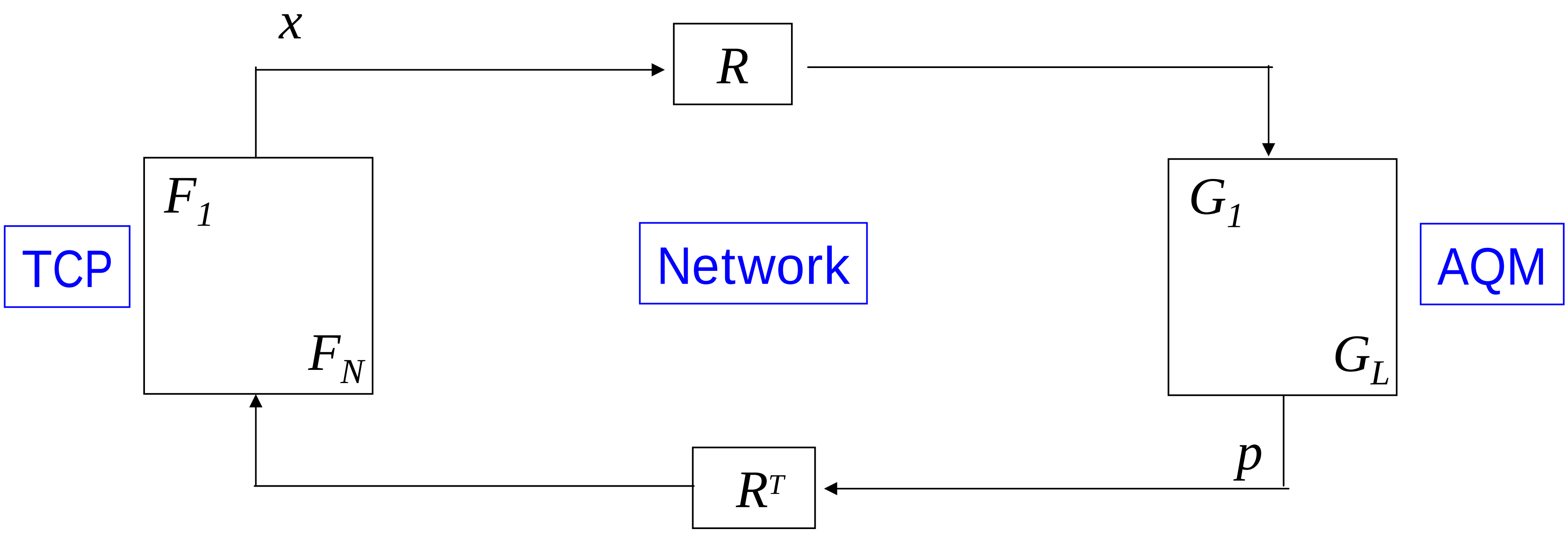
$$F_i = \frac{1}{T_i^2} - \frac{x_i^2}{2} \sum_l R_{li} p_l \longleftarrow \dot{p}_l = g_l\left(p_l(t), \sum_i R_{li} x_i(t)\right)$$

# Homogeneous protocol




$$x_i(t+1) \;=\; F_i\left(\sum_l R_{li}\, p_l(t),\ x_i(t)\right)$$

# Heterogeneous protocol




$$x_i(t+1) \;=\; F_i\left(\sum_l R_{li}\, p_l(t),\; x_i(t)\right)$$

$$x_i^j(t+1) \;=\; F_i^j\left(\sum_l R_{li}\, m_l^j\big(p_l(t)\big),\; x_i^j(t)\right)$$

heterogeneous prices for type *j* sources

# Heterogeneous protocols

- Equilibrium: $p$ that satisfies

$$x_i^j(p) \;=\; f_i^j\left(\sum_l R_{li} m_l^j(p_l)\right)$$

$$y_l(p) := \sum_{i,j} R_{li}^j x_i^j(p) \begin{cases} \leq c_l \\ = c_l & \text{if } \; p_l > 0 \end{cases}$$

Duality model no longer applies !

- $p_l$ can no longer serve as Lagrange multiplier

# Heterogeneous protocols

- Equilibrium: $p$ that satisfies

$$x_i^j(p) = f_i^j\left(\sum_l R_{li} m_l^j(p_l)\right)$$

$$y_l(p) := \sum_{i,j} R_{li}^j x_i^j(p) \begin{cases} \leq c_l \\ = c_l & \text{if } p_l > 0 \end{cases}$$

Need to re-examine all issues

- Equilibrium: exists? unique? efficient? fair?
- Dynamics: stable? limit cycle? chaotic?
- Practical networks: typical behavior? design guidelines?

# Heterogeneous protocols

- Equilibrium: $p$ that satisfies

$$x_i^j(p) = f_i^j\left(\sum_l R_{li} m_l^j(p_l)\right)$$

$$y_l(p) := \sum_{i,j} R_{li}^j x_i^j(p) \begin{cases} \le c_l \\ = c_l & \text{if } p_l > 0 \end{cases}$$

- Dynamic: dual algorithm

$$x_i^j(p(t)) = f_i^j\left(\sum_l R_{li} m_l^j(p_l(t))\right)$$

$$\dot{p}_l = \gamma_l\left(y_l(p(t)) - c_l\right)$$

# Notation

- Simpler notation: $p$ is *equilibrium* if

  $y(p) \; = \; c$ on bottleneck links

- Jacobian: $\mathbf{J}(p) := \dfrac{\partial y}{\partial p}(p)$

- Linearized dual algorithm:

  $$\partial \dot{p} = \gamma \, \mathbf{J}(p^*) \; \partial p(t)$$

See Simsek, Ozdaglar, Acemoglu 2005 for generalization

# Outline

- Review: homogeneous case
- Motivating experiments
- Model
- Equilibrium
  - Existence, uniqueness, local stability
  - Efficiency, fairness
- Slow timescale control

# Existence

**Theorem**

Equilibrium $p$ exists, despite lack of underlying utility maximization

- Generally non-unique
  - There are networks with unique bottleneck set but infinitely many equilibria
  - There are networks with multiple bottleneck set each with a unique (but distinct) equilibrium

# Regular networks

**Definition**

A *regular network* is a tuple $(R, c, m, U)$ for which all equilibria $p$ are locally unique, i.e.,

$$\det \mathbf{J}(p) := \det \frac{\partial y}{\partial p}(p) \neq 0$$

**Theorem**

- Almost all networks are regular
- A regular network has finitely many and odd number of equilibria (e.g. 1)

# Regular networks

Proof idea:

- Sard's Theorem: critical value of a continuously differentiable function over open set has measure zero
- Apply to $y(p) = c$ on each bottleneck set ➔ regularity
- Compact equilibrium set ➔ finiteness

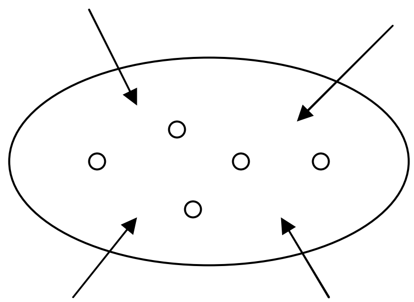
# Regular networks

$$\text{index } I(p) := \begin{cases} -1 & \text{if } \det \mathbf{J}(p) < 0 \\ 1 & \text{if } \det \mathbf{J}(p) > 0 \end{cases}$$

<u>Proof idea:</u>

- Poincare Hopf index theorem: if there exists a vector field with non-singular $dv/dp$ at every equilibrium and all trajectories move inward, then

$$\sum_{\text{eq } p} I(p) = (-1)^L$$



- Dual algorithm defines such a vector field
- Index theorem implies odd #equilibria

# Global uniqueness

$$\dot{m}_l^j \in [a_l, 2^{1/L} a_l] \quad \text{for any } a_l > 0$$
$$\dot{m}_l^j \in [a^j, 2^{1/L} a^j] \quad \text{for any } a^j > 0$$

**Theorem**

- If *price heterogeneity* is small, then equilibrium is globally unique

**Corollary**

- If price mapping functions $m_l^j$ are linear and link independent, then equilibrium is globally unique

e.g. a network of RED routers with slope inversely proportional to link capacity almost always has globally unique equilibrium

# Global uniqueness

$$\dot{m}_l^j \in [a_l, 2^{1/L} a_l] \quad \text{for any } a_l > 0$$

$$\dot{m}_l^j \in [a^j, 2^{1/L} a^j] \quad \text{for any } a^j > 0$$

**Theorem**

- If *price heterogeneity* is small, then equilibrium is globally unique

Remarks:

- Condition independent of $U, R, c$
- Depends on $m$ and size $L$ of network
- "Tight" from Index Theorem

# Local stability: `uniqueness' → stability

$$\dot{m}_l^j \in [a_l, 2^{1/L} a_l] \quad \text{for any } a_l > 0$$

$$\dot{m}_l^j \in [a^j, 2^{1/L} a^j] \quad \text{for any } a^j > 0$$

**Theorem**

- If *price heterogeneity* is small, then the unique equilibrium $p$ is locally stable

Linearized dual algorithm: $\delta\dot{p} = \gamma\, \mathbf{J}(p^*)\ \delta p(t)$

Equilibrium $p$ is *locally stable* if

$$\mathrm{Re}\,\lambda\big(\mathbf{J}(p)\big) < 0$$

# Local stability: `converse'

**Theorem**

- If all equilibria $p$ are locally stable, then it is globally unique

Proof idea:

- For all equilibrium $p$: $I(p) = (-1)^L$
- Index theorem:

$$\sum_{\text{eq } p} I(p) = (-1)^L$$

# Outline

- Review: homogeneous case
- Motivating experiments
- Model
- Equilibrium
  - Existence, uniqueness, local stability
  - Efficiency, fairness
- Slow timescale control

Tang, Wang, Low, Chiang. Infocom 2005

# Efficiency

**Theorem**

- Every equilibrium $p^*$ is Pareto efficient

Proof:

- Every equilibrium $p^*$ yields a (unique) rate $x(p^*)$ that solves

$$\max_{x \geq 0} \sum_j \sum_i \lambda_i^j(p^*) U_i^j(x_i^j) \quad \text{s.t.} \quad Rx \leq c$$

# Efficiency

**Theorem**

- Every equilibrium $p^*$ is Pareto efficient

- Measure of optimality

$$V^* \; := \; \max_{x \geq 0} \sum_j \sum_i U_i^j(x_i^j) \quad \text{s.t.} \quad Rx \leq c$$

- Achieved: $V(p^*) \; := \; \sum_j \sum_i U_i^j(x_i^j(p^*))$

# Efficiency

**Theorem**

- Every equilibrium $p^*$ is Pareto efficient
- Loss of optimality:

$$\frac{V(p^*)}{V^*} \geq \frac{\min \ \dot{m}_l^j}{\max \ \dot{m}_l^j}$$

- Measure of optimality

$$V^* \ := \ \max_{x \geq 0} \sum_j \sum_i U_i^j(x_i^j) \quad \text{s.t.} \quad Rx \leq c$$

- Achieved: $\quad V(p^*) \ := \ \sum_j \sum_i U_i^j(x_i^j(p^*))$

# Efficiency

**Theorem**

- Every equilibrium $p^*$ is Pareto efficient
- Loss of optimality:

$$\frac{V(p^*)}{V^*} \geq \frac{\min \ \dot{m}_l^j}{\max \ \dot{m}_l^j}$$

e.g. A network of RED routers with default parameters suffers no loss of optimality

# Intra-protocol fairness

**Theorem**

- Fairness among flows within each type is unaffected, i.e., still determined by their utility functions and Kelly's problem with reduced link capacities

Proof idea:

- Each equilibrium $p$ chooses a partition of link capacities among types, $c^j := c^j(p)$
- Rates $x^j(p)$ then solve

$$\max_{x^j \geq 0} \sum_i U_i^j(x_i^j) \quad \text{s.t.} \quad R^j x^j \leq c^j$$

# Inter-protocol fairness

**Theorem**

- Any fairness is achievable with a linear scaling of utility functions

$$\bar{x}^j \; := \; \arg\max_{x^j \geq 0} \sum_i U_i^j(x_i^j) \quad \text{s.t.} \quad R^j x^j \leq c$$

$$\text{all achievable rates} \;\; X \; := \left\{ x = \sum_j a^j \bar{x}^j \right\}$$

# Inter-protocol fairness

**Theorem**

- Any fairness is achievable with a linear scaling of utility functions
- i.e. given any $x$ in $X$, there exists $\mu$ s.t. an equilibrium $p$ has $x(p) = x$

$$x_i^j(p(t)) = f_i^j\left(\frac{1}{\mu_i^j}\sum_l R_{li} m_l^j(p_l(t))\right)$$

$$\dot{p}_l = \gamma_l\left(y_l(p(t)) - c_l\right)$$

# Outline

- Review: homogeneous case
- Motivating experiments
- Model
- Equilibrium
  - Existence, uniqueness, local stability
  - Efficiency, fairness
- Slow timescale control

# Slow timescale control

Slow timescale scaling of utility function

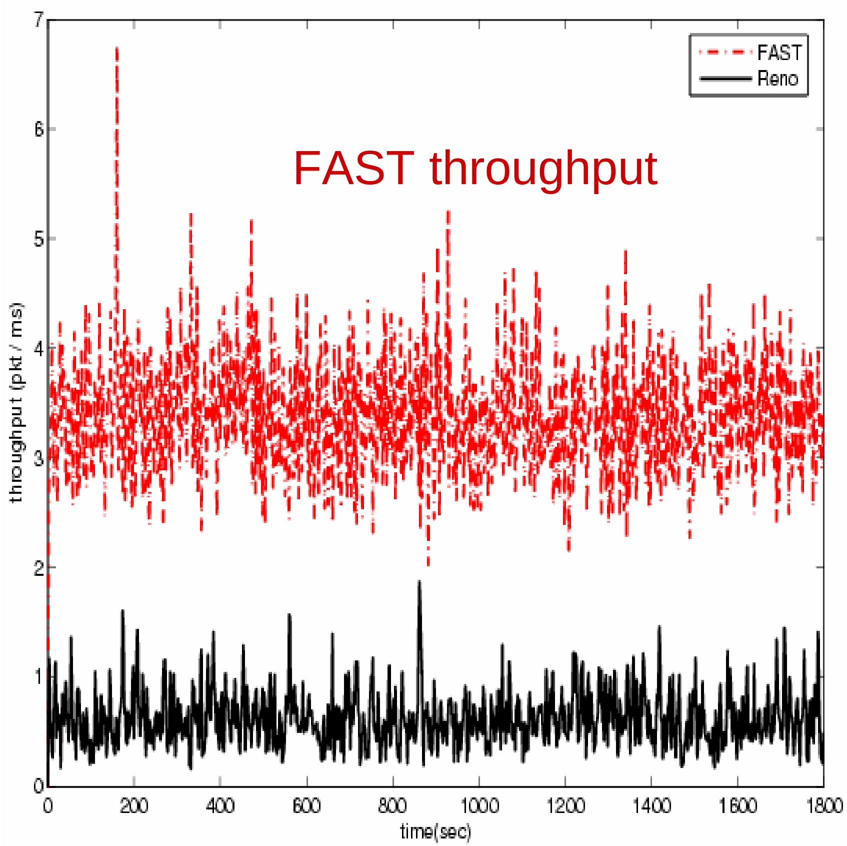
$$x_i^j(t) \;=\; f_i^j\left(\frac{q_i^j(t)}{\mu_i^j(t)}\right)$$

scaling of end--to-end price

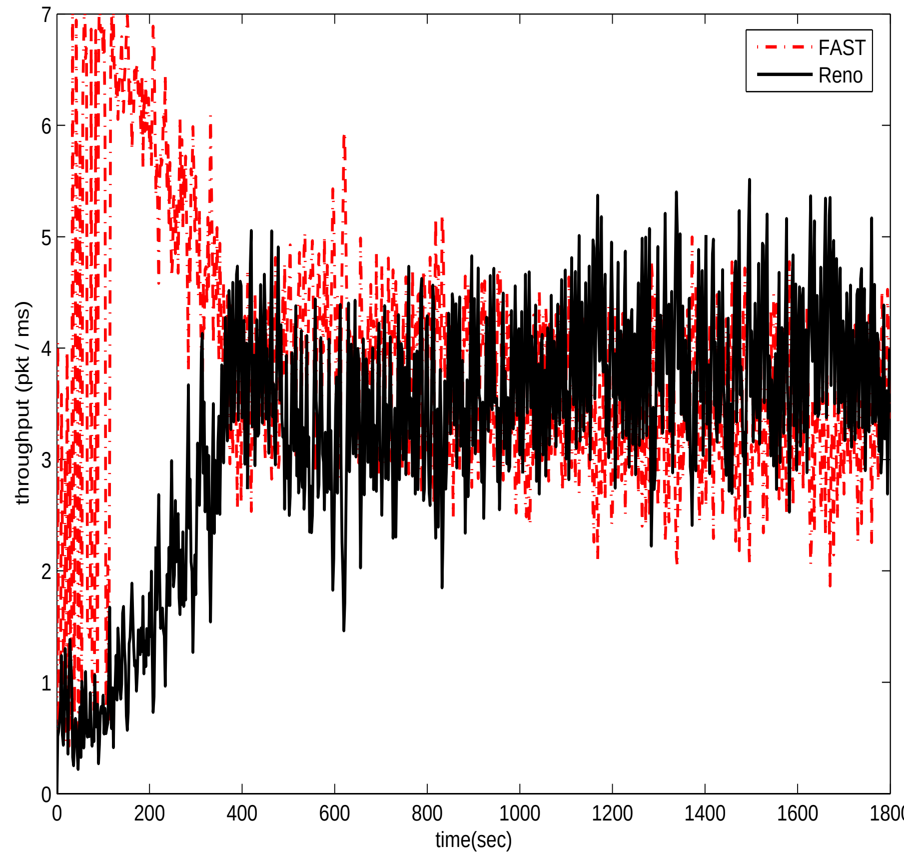
$$\mu_i^j(\mathrm{t}+1) \;=\; \kappa_i^j \mu_i^j(t) + (1-\kappa_i^j)\,\frac{\sum_l m_l^j(p_l(t))}{\sum_l p_l(t)}$$

slow timescale update of scaling factor

# ns2 simulation: buffer=80pks

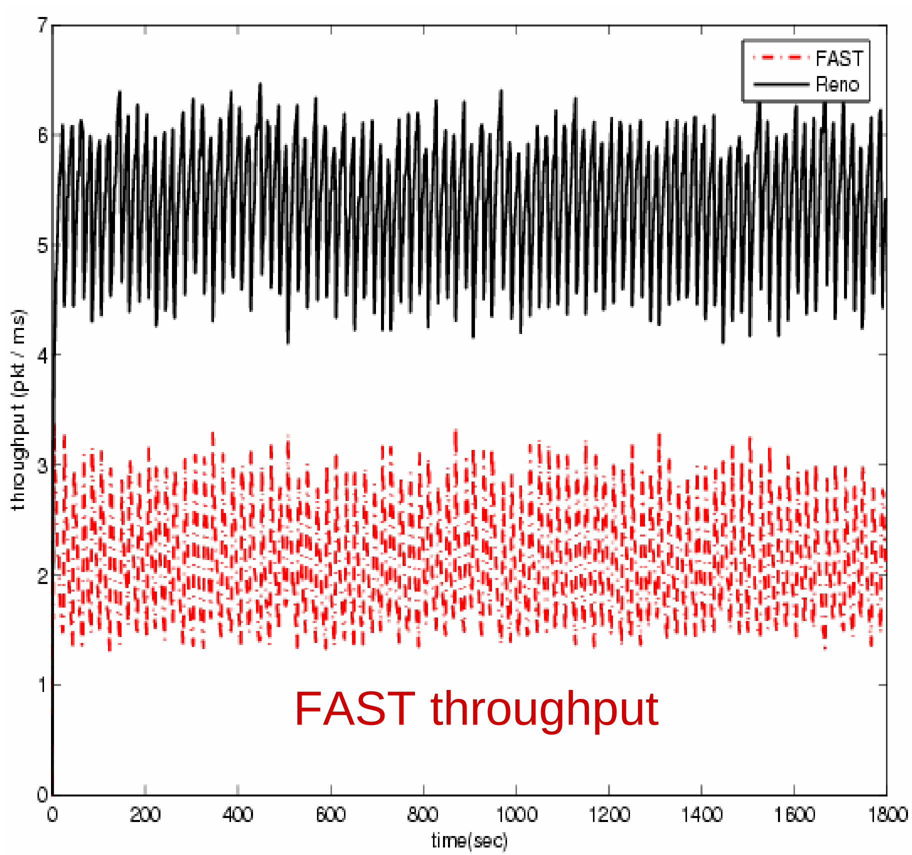



without slow timescale control

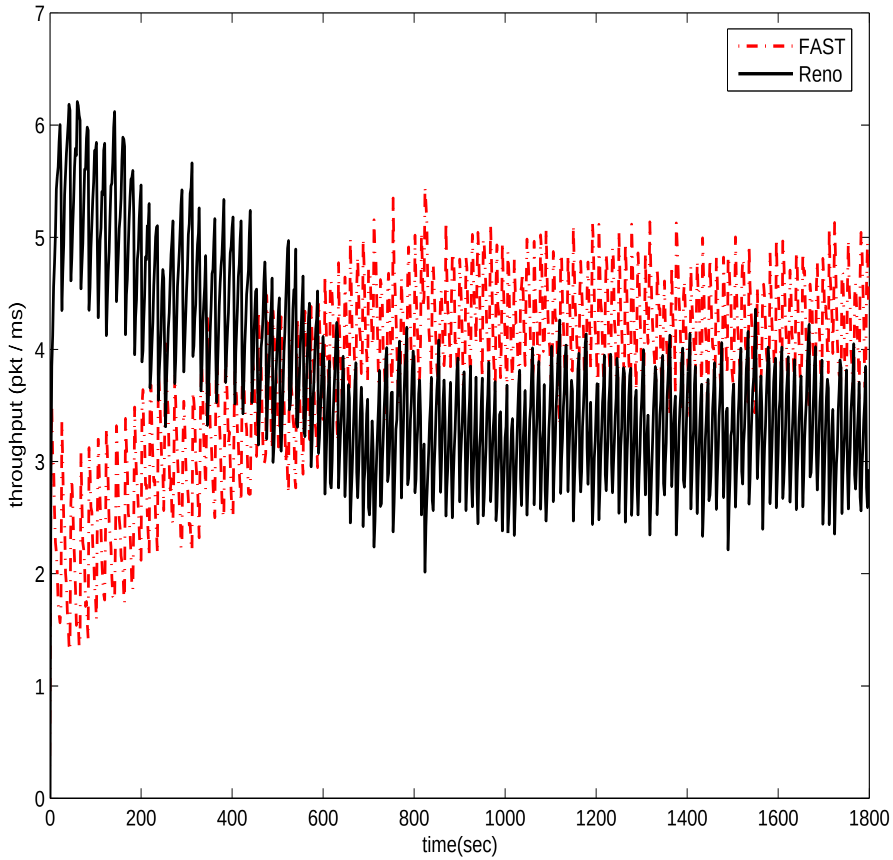



with slow timescale control

# ns2 simulation: buffer=400pks




without slow timescale control




with slow timescale control

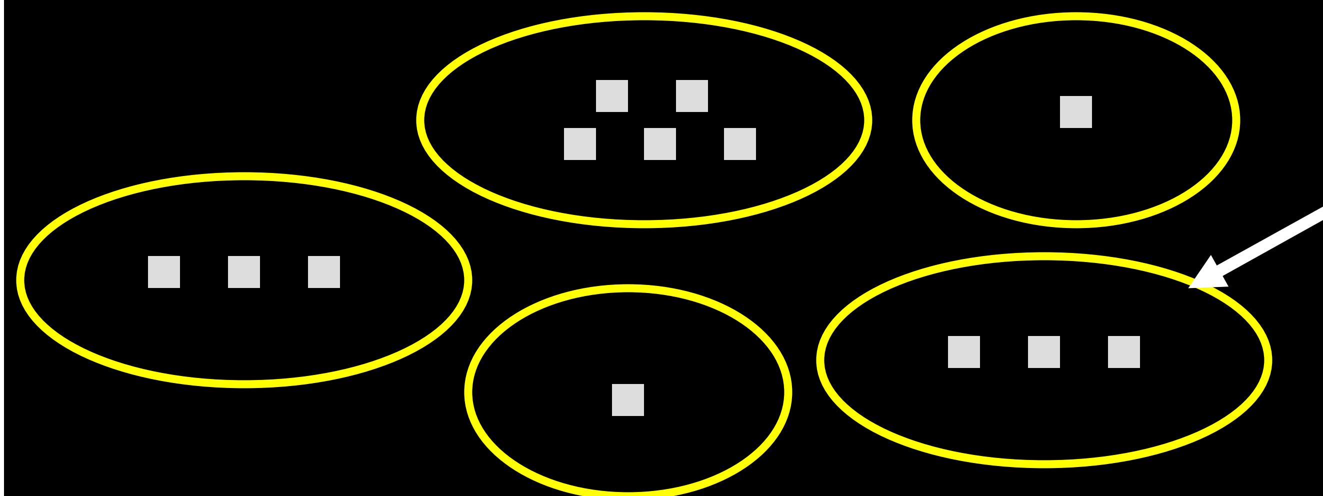
# Summary: equilibrium structure

**Uni-protocol**

- Unique bottleneck set
- Unique rates & prices

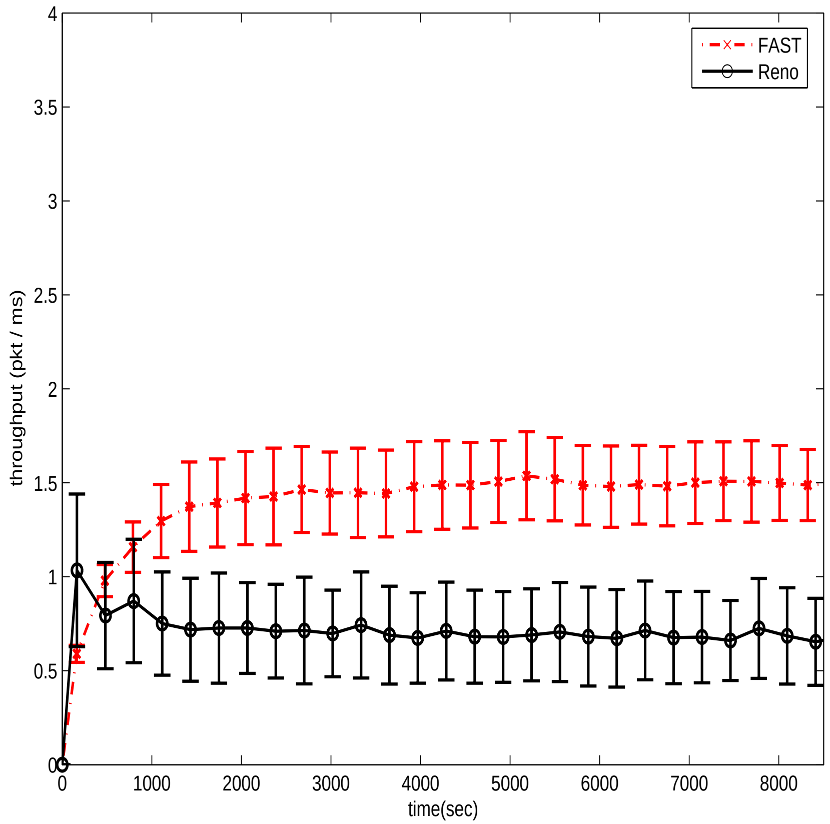
**Multi-protocol**

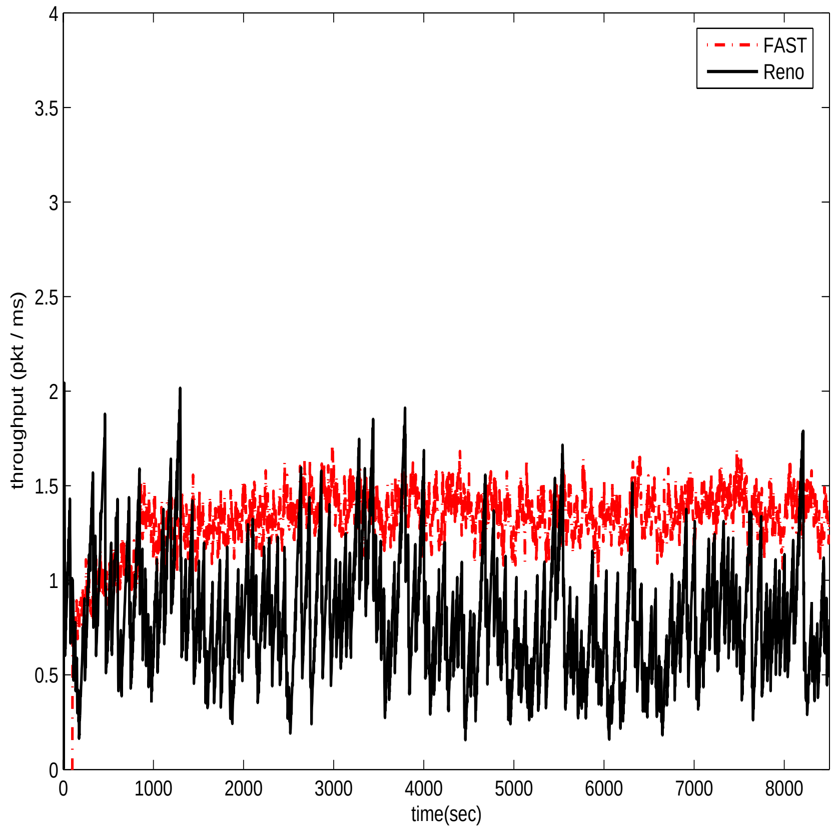
- Non-unique bottleneck sets
- Non-unique rates & prices for each B.S.



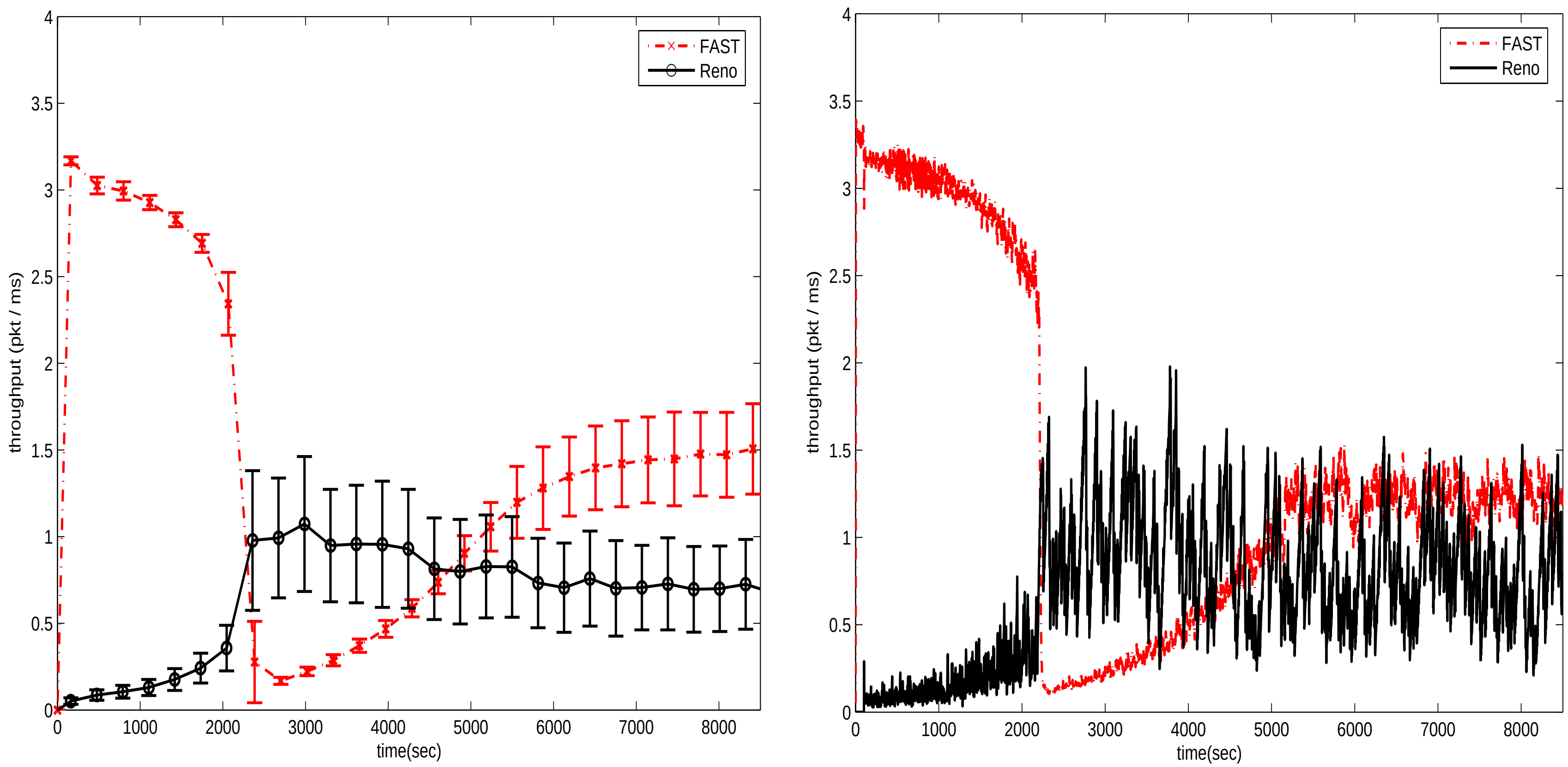
- always odd
- not all stable
- uniqueness conditions

# Reno First


FAST
Reno
throughput (pkt / ms)
time(sec)
4
3.5
3
2.5
2
1.5
1
0.5
0
0
1000
2000
3000
4000
5000
6000
7000
8000


FAST
Reno
throughput (pkt / ms)
time(sec)
4
3.5
3
2.5
2
1.5
1
0.5
0
0
1000
2000
3000
4000
5000
6000
7000
8000

# FAST First



FAST
Reno
throughput (pkt / ms)
time(sec)
0
0.5
1
1.5
2
2.5
3
3.5
4
1000
2000
3000
4000
5000
6000
7000
8000
FAST
Reno
throughput (pkt / ms)
time(sec)

# Alpha trajectory