

# TRANSFORM METHODS FOR HEAVY TRAFFIC ANALYSIS

---

Daniela Hurtado Lange, Siva Theja Maguluri

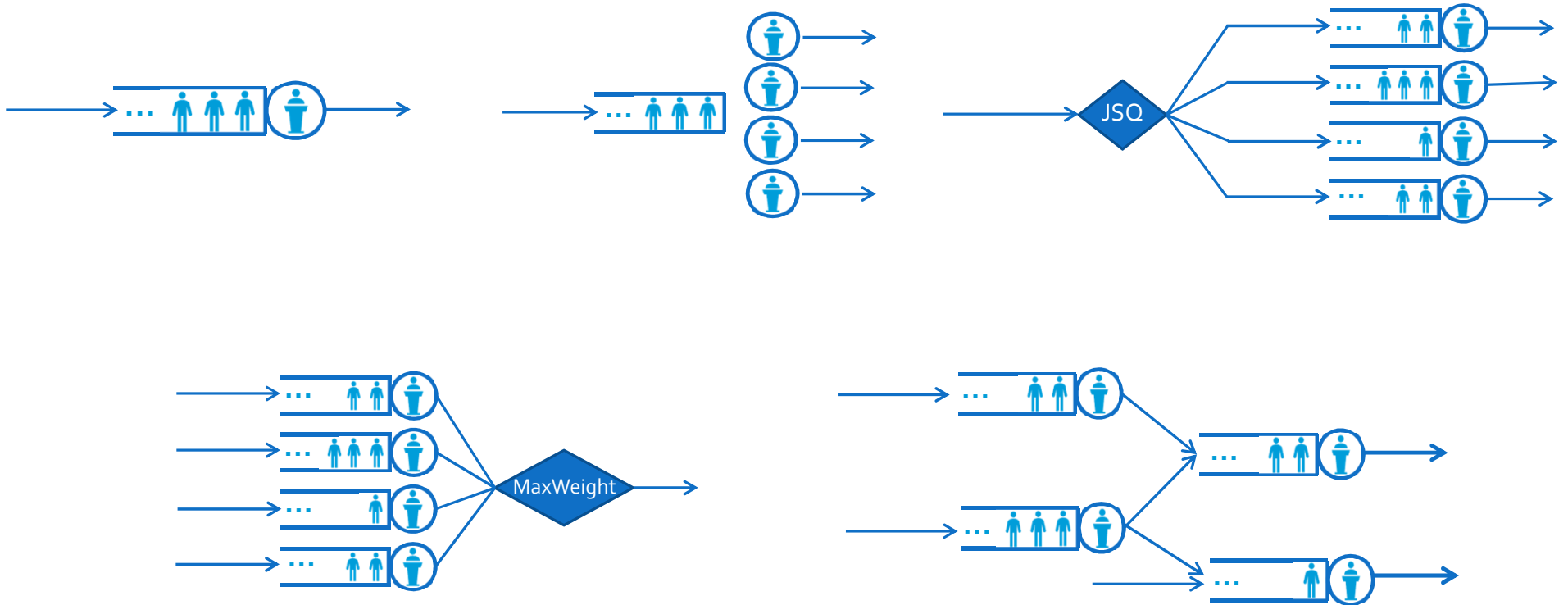
Industrial and Systems Engineering

Georgia Institute of Technology

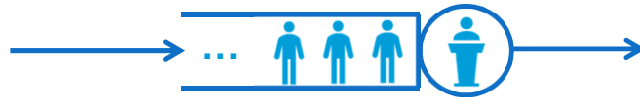
# OUTLINE

- Motivation
- Transform techniques
  - Single server queue
  - Load balancing system
- Conclusion and future work

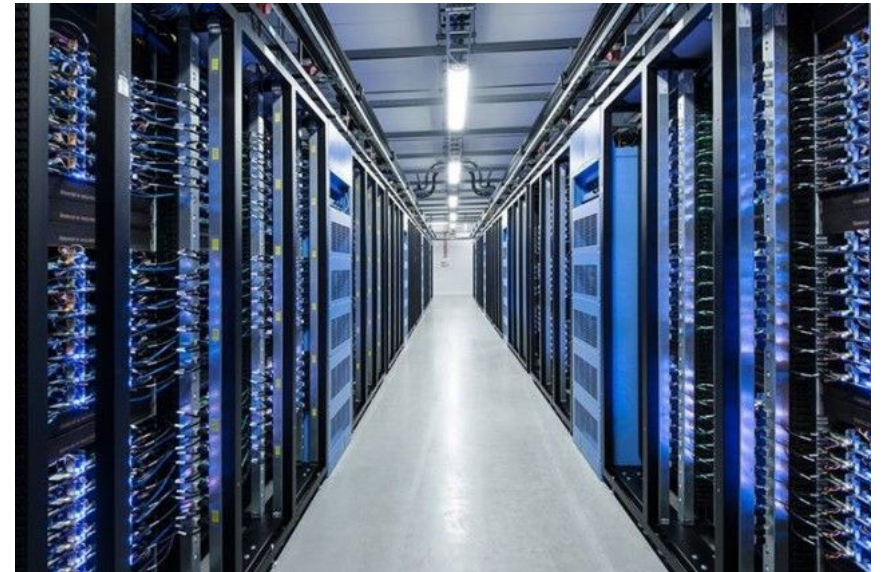
# QUEUEING SYSTEMS



# QUEUEING ASYMPTOTICS



Heavy Traffic



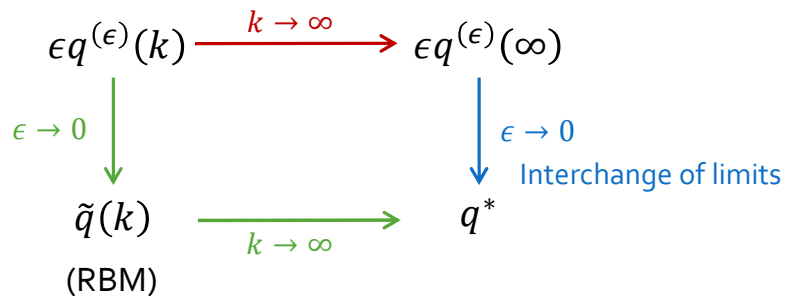
Mean Field

# HEAVY-TRAFFIC

- Load queueing system close to maximum capacity

## DIFFUSION LIMITS

- Scale space and time:  $\epsilon = 1/\sqrt{n}$



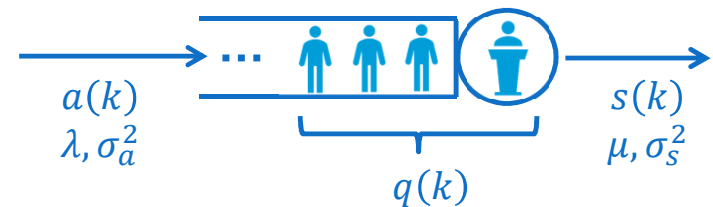
## DRIFT METHOD [Eryilmaz, Srikant 13]

- Set to zero the drift of a test function
- $k \rightarrow \infty$ :
 
$$E \left[ V \left( q^{(\epsilon)}(k) \right) \right] = E \left[ V \left( q^{(\epsilon)}(k + 1) \right) \right]$$
- $\epsilon \rightarrow 0$ :  $E[V'(q^*)]$
- No interchange of limits
  - Direct method
- How to choose  $V(\cdot)$ ?

# SINGLE SERVER QUEUE

# SINGLE SERVER QUEUE MODEL

- Discrete time model
- $q(k)$ : # jobs in the system at the beginning of time slot  $k$
- $a(k)$ : # arrivals in time slot  $k$ 
  - Mean  $\lambda$  and variance  $\sigma_a^2$
- $s(k)$ : offered service in time slot  $k$ 
  - Mean  $\mu$  and variance  $\sigma_s^2$
- Both are i.i.d. sequences, independent of each other

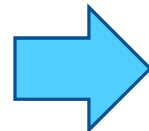


Assume positive recurrence  
 $\Leftrightarrow \epsilon = \mu - \lambda > 0$

- Dynamics of the queues

$$q(k + 1) = [q(k) + a(k) - s(k)]^+$$

$$= q(k) + a(k) - s(k) + u(k)$$



$$q(k + 1)u(k) = 0$$

Unused service

# HEAVY TRAFFIC – DRIFT METHOD

- Kingman, 1962:

- In steady state:

$$E[q^2(k+1)] = E[q^2(k)]$$

- Yields:

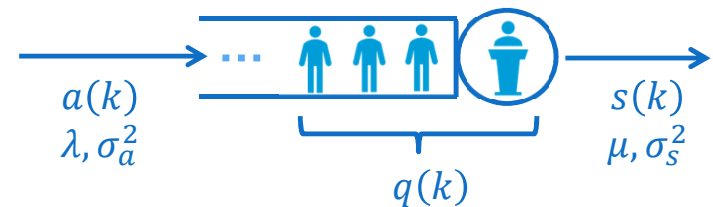
$$E[q] = \frac{E[(a-s)^2]}{2(\mu-\lambda)} + \frac{E[u^2]}{2(\mu-\lambda)}$$

$\underbrace{\hspace{10em}}_{\epsilon}$

It can be shown that this term is at most a constant and so small compared to the first term when  $\lambda \rightarrow \mu$

- Therefore,

$$\lim_{\epsilon \downarrow 0} E[\epsilon q] = \frac{\sigma_a^2 + \sigma_s^2}{2}$$



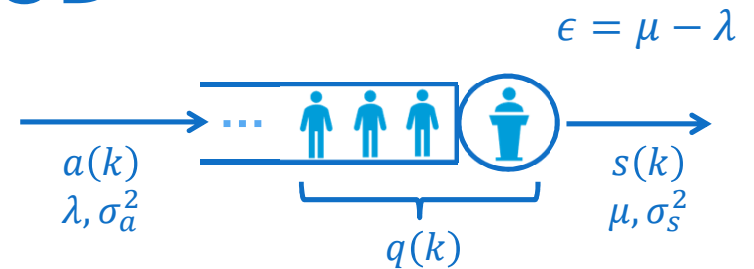
$$q(k+1) = [q(k) + a(k) - s(k)]^+ = q(k) + a(k) - s(k) + u(k)$$



$$q(k+1)u(k) = 0$$

# DRIFT METHOD

- Drift method [Eryilmaz, Srikant 13]:



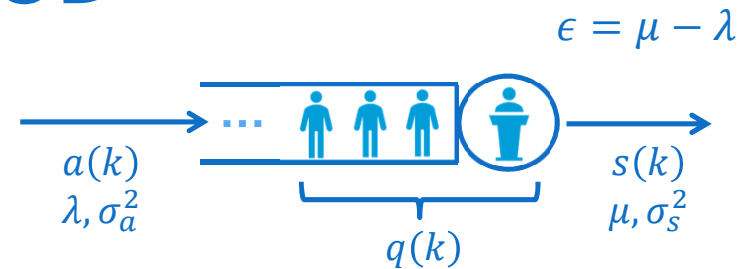
$V(q)$ :

$$\begin{array}{cccc}
 & q^2 & q^3 & \dots & q^{m+1} \\
 & \downarrow & \downarrow & & \downarrow \\
 \text{Obtain:} & E[\epsilon q] & E[\epsilon^2 q^2] & \dots & E[\epsilon^m q^m] \\
 & \downarrow & \downarrow & & \downarrow \\
 & \frac{\sigma_a^2 + \sigma_s^2}{2} & 2 \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)^2 & & m! \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)^m
 \end{array}$$

$$\epsilon q \Rightarrow \text{Exp}\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)$$

# DRIFT METHOD

- Drift method [Eryilmaz, Srikant 13]:



$$V(q): \quad 1 + \epsilon q + \frac{1}{2} \epsilon^2 q^2 + \frac{1}{3!} \epsilon^3 q^3 + \dots + \frac{1}{(m+1)!} \epsilon^{m+1} q^{m+1} + \dots = e^{\epsilon q}$$

$$\text{Obtain:} \quad 1 + E[\epsilon q] + \frac{1}{2} E[\epsilon^2 q^2] + \dots + \frac{1}{m!} E[\epsilon^m q^m] + \dots = E[e^{\epsilon q}]$$

$$\downarrow$$

$$\frac{\sigma_a^2 + \sigma_s^2}{2}$$

$$\downarrow$$

$$2 \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)^2$$

$$\downarrow$$

$$m! \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)^m$$

$$\epsilon q \Rightarrow \text{Exp} \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)$$

**IDEA**

Use  $e^{\theta \epsilon q}$  as a test function

# MGF METHOD [Hurtado-Lange, Maguluri 19]

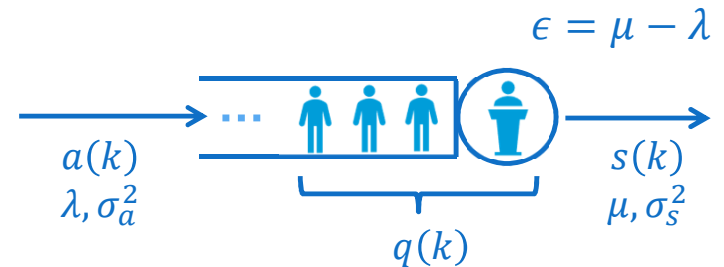
- In steady-state:  $E[e^{\theta\epsilon q(k+1)}] = E[e^{\theta\epsilon q(k)}]$

- Key Lemma:

$$(e^{\theta\epsilon q(k+1)} - 1)(e^{-\theta\epsilon u(k)} - 1) = 0$$

- From the lemma:

$$E[e^{\theta\epsilon q(k+1)}] = E[e^{\theta\epsilon(q(k)+a(k)-s(k))} - e^{-\theta\epsilon u(k)} + 1]$$



$$q(k+1) = q(k) + a(k) - s(k) + u(k)$$

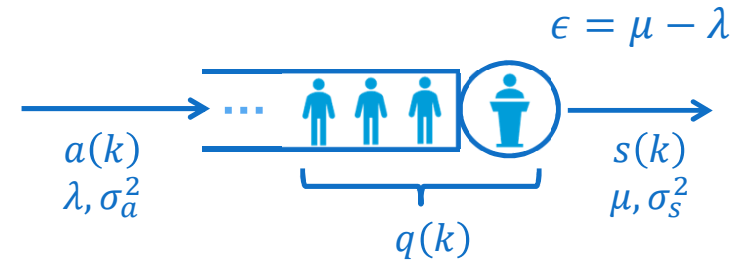


$$q(k+1)u(k) = 0$$

# MGF METHOD

$$\begin{aligned} \Rightarrow E[e^{\theta\epsilon q}] &= \frac{1 - E[e^{-\theta\epsilon u}]}{1 - E[e^{\theta\epsilon(a-s)}]} \\ &= \frac{\theta\epsilon E[u] + h.o.t.}{-\theta\epsilon E[a-s] - \frac{\theta^2\epsilon^2}{2} E[(a-s)^2] + h.o.t.} \\ &= \frac{\theta\epsilon^2 + h.o.t.}{\theta\epsilon^2 - \frac{\theta^2\epsilon^2}{2}(\sigma_a^2 + \sigma_s^2 + \epsilon^2) + h.o.t.} \end{aligned}$$

$\theta/2$  (pointing to  $\frac{\theta^2\epsilon^2}{2}$ )  
 $1$  (pointing to  $\theta\epsilon^2$ )



$$(e^{\theta\epsilon q(k+1)} - 1)(e^{-\theta\epsilon u(k)} - 1) = 0$$



$$E[e^{\theta\epsilon q(k)}] = E[e^{\theta\epsilon(q(k)+a(k)-s(k))} - e^{-\theta\epsilon u(k)} + 1]$$

$\epsilon \downarrow 0$  →

$$\frac{1}{1 - \theta \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)}$$

MGF of expo. r.v.  
with mean  $\frac{\sigma_a^2 + \sigma_s^2}{2}$

$$\epsilon q \Rightarrow \text{Expo} \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)$$

# TRANSFORM TECHNIQUES

**Step 1** Show  $q^{(\epsilon)}(k)$  is positive recurrent. Then,  $q^{(\epsilon)}(k) \xrightarrow{k \rightarrow \infty} q^{(\epsilon)}(\infty)$   
**Step 2** Set Drift of an exponential function to zero under the stationary distribution

## Moment Generating Function $E[e^{\theta \epsilon q}]$

- Two-sided Laplace transform of stationary distribution
- Must exist for  $\theta$  in an open interval around origin  
→ Must prove

## Characteristic Function $E[e^{j\omega \epsilon q}]$

- Fourier Transform
- Always exists
- Complex numbers
- Existence of moments  
→ Must prove

## One sided Laplace Transform

- $E[e^{\theta \epsilon q}]$  for  $\theta < 0$
- Always exists, because  $q \geq 0$
- Existence of moments  
→ Must prove
- Related to BAR approach  
[Braverman, Dai, Miyazawa 17]

# RATES OF CONVERGENCE

- Drift method gives

$$\frac{1}{\epsilon} \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right) - B(\epsilon) \leq E[q] \leq \frac{1}{\epsilon} \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right) + B(\epsilon)$$

- True for all  $\epsilon$
- Rate of convergence of the mean
- Rate of convergence of the distribution
  - Ongoing work

# KEY TAKEAWAYS

- Set the drift of a function to zero
  - Drift Method – Quadratic and Polynomial test functions
  - Transform Techniques – Exponential test function
- Unused service is important

$$q(k + 1)u(k) = 0$$

$$(e^{\theta \epsilon q(k+1)} - 1)(e^{-\theta \epsilon u(k)} - 1) = 0$$

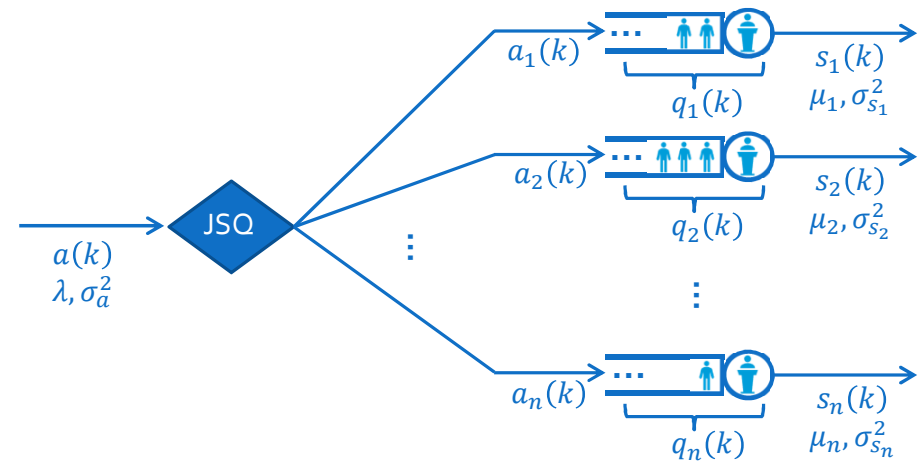
# LOAD BALANCING SYSTEM

---

Complete Resource Pooling

# LOAD BALANCING SYSTEM

- Discrete time model
- $n$  servers, with their own queue
- One stream of i.i.d. arrivals
  - Routing: Join the Shortest Queue (JSQ) policy
- Heavy-Traffic Analysis [Foschini, Salz '78]
  - Diffusion Limit Analysis



$$\begin{aligned}
 q_i(k+1) &= [q_i(k) + a_i(k) - s_i(k)]^+ \\
 &= q_i(k) + a_i(k) - s_i(k) + u_i(k)
 \end{aligned}$$



$$q_i(k+1)u_i(k) = 0 \text{ but } q_i(k+1)u_j(k) \neq 0 \text{ if } i \neq j$$

# STATE SPACE COLLAPSE

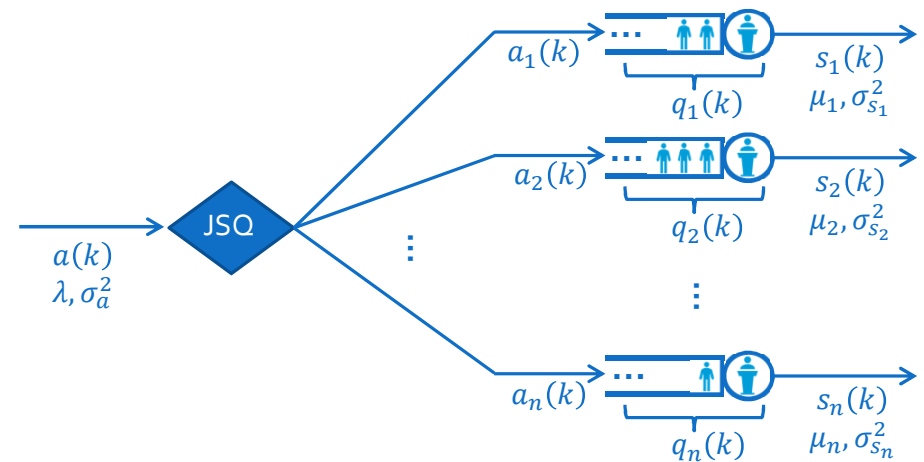
- Heavy-traffic:

$$\epsilon = \sum_i \mu_i - \lambda > 0$$

- SSC[Eryilmaz, Srikant 13]:

All queues are (approximately) equal in the heavy-traffic limit

$$q_i \approx \frac{\sum_j q_j}{n} \quad \forall i$$



$$\mathbf{q}(k+1) = \mathbf{q}(k) + \mathbf{a}(k) - \mathbf{s}(k) + \mathbf{u}(k)$$

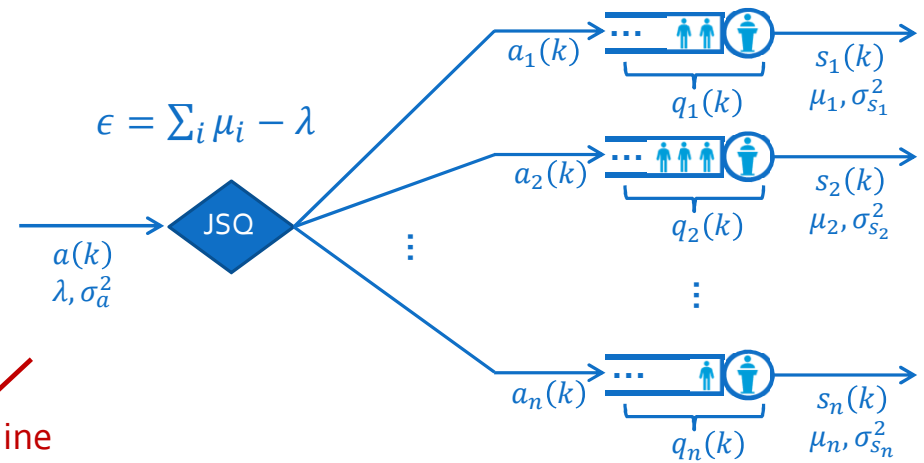
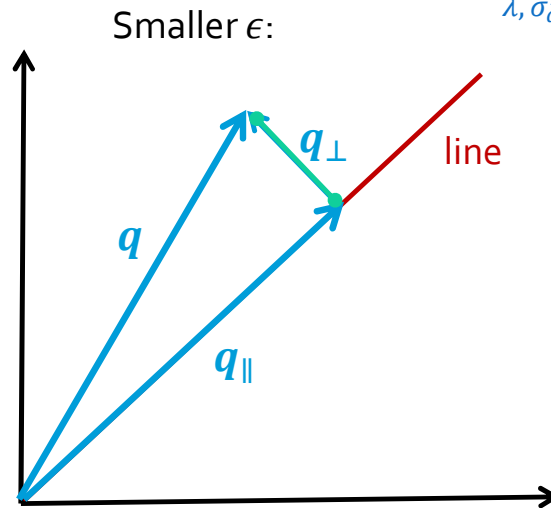
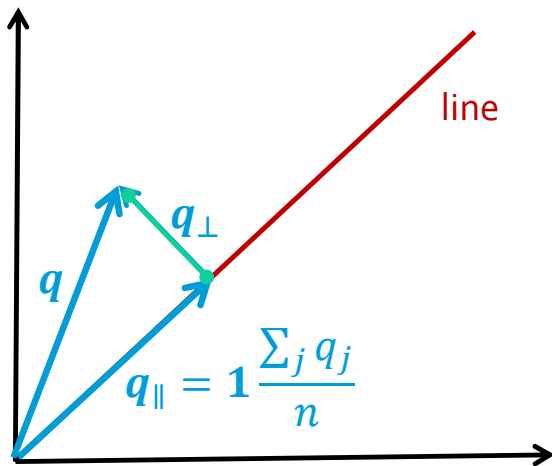
# SSC INTERPRETATION

**Proposition**[Eryilmaz, Srikant 13]:

$\mathbf{q}_{\parallel}(k)$ : projection of  $\mathbf{q}(k)$  on the line

$\mathbf{q}_{\perp}(k) := \mathbf{q}(k) - \mathbf{q}_{\parallel}(k)$  : error of approximating  $\mathbf{q} \approx \mathbf{q}_{\parallel}$

Then,  $E[\|\mathbf{q}_{\perp}\|^m] \leq M_m$  for all  $m = 1, 2, \dots$



$\mathbf{q} \approx \mathbf{q}_{\parallel}$  as  $\epsilon \downarrow 0$

# UNUSED SERVICE

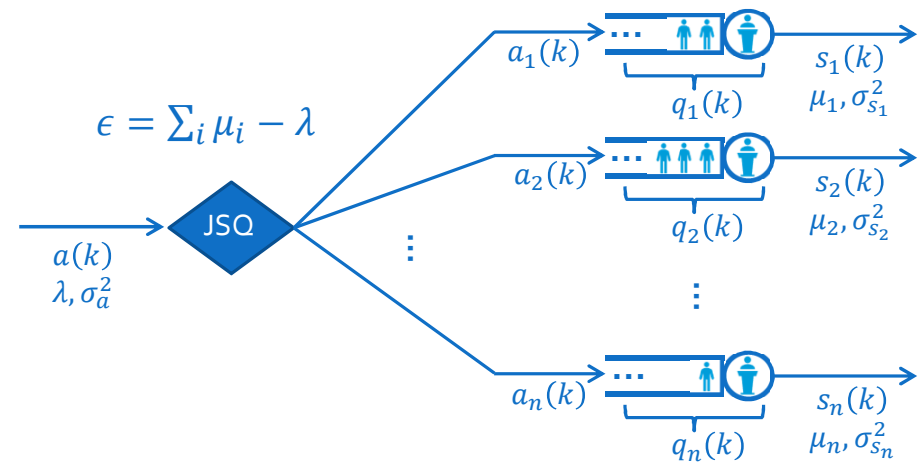
- Recall in a single server queue  $q(k+1)u(k) = 0$

- Under SSC,  $(\sum_i q_i(k+1))(\sum_i u_i(k)) \approx 0$

- $q_i(k+1)u_i(k) = 0$

- $q_i(k+1)u_j(k) \approx q_j(k+1)u_j(k) = 0$

- $\sum_i q_i(k)$  behaves like a single queue

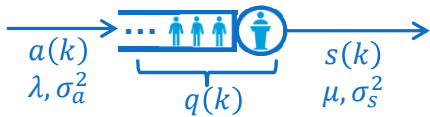


$$\mathbf{q}(k+1) = \mathbf{q}(k) + \mathbf{a}(k) - \mathbf{s}(k) + \mathbf{u}(k)$$

MGF method: Set to zero the drift of  $V(\mathbf{q}) = e^{\theta \epsilon \sum_i q_i}$  [Hurtado-Lange, M '19]

$$E\left[\left(e^{\theta \epsilon \sum_i q_i(k+1)} - 1\right)\left(e^{-\theta \epsilon \sum_i u_i(k)} - 1\right)\right] \text{ is } o(\epsilon^2)$$

# MGF METHOD [Hurtado Lange, Maguluri 19]



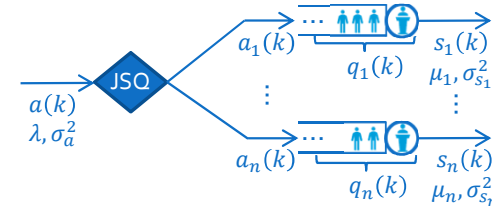
- Set drift of  $V(q) = e^{\theta \epsilon q}$  to zero

$$(e^{\theta \epsilon q(k+1)} - 1)(e^{-\theta \epsilon u(k)} - 1) = 0$$

- Yields

$$\lim_{\epsilon \rightarrow 0} E[e^{\theta \epsilon q}] = \frac{1}{1 - \theta \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)}$$

$$\epsilon q \Rightarrow \text{Expo} \left( \frac{\sigma_a^2 + \sigma_s^2}{2} \right)$$



$$\mathbf{q}_{\parallel} = \mathbf{1} \frac{\sum_i q_i}{n}$$

$$\& \mathbf{q}_{\perp} = \mathbf{q} - \mathbf{q}_{\parallel}$$

- Set drift of  $V(\mathbf{q}) = e^{\theta \epsilon \sum_i q_i}$  to zero

$$E[(e^{\theta \epsilon \sum_i q_i(k+1)} - 1)(e^{-\theta \epsilon \sum_i u_i(k)} - 1)] \text{ is } o(\epsilon^2)$$

- Yields

$$\lim_{\epsilon \rightarrow 0} E[e^{\theta \epsilon \sum_i q_i}] = \frac{1}{1 - \theta \left( \frac{\sigma_a^2 + \sum_i \sigma_{s_i}^2}{2} \right)}$$

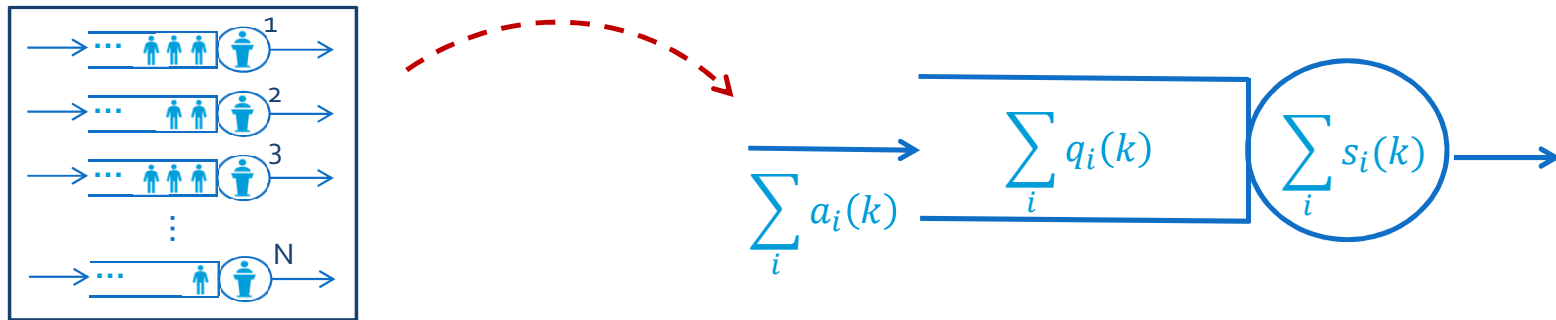
- Then,  $\epsilon \mathbf{q}_{\parallel} \xrightarrow{d} \frac{1}{n} \text{Expo} \left( \frac{\sigma_a^2 + \sum_i \sigma_{s_i}^2}{2} \right)$

- SSC implies  $\epsilon \mathbf{q}_{\perp} \Rightarrow \mathbf{0}$

$$\epsilon \mathbf{q} \Rightarrow \frac{1}{n} \text{Expo} \left( \frac{\sigma_a^2 + \sum_i \sigma_{s_i}^2}{2} \right)$$

# COMPLETE RESOURCE POOLING

- Complete Resource Pooling  $\leftrightarrow$  One dimensional State Space Collapse



- Many systems exhibit this behavior
  - Ad hoc wireless network
  - Generalized Switch Model
- We can use the MGF method in all these queueing systems!

# KEY TAKEAWAYS

- Complete Resource Pooling
  - Single bottle-neck resource
  - One dimensional State Space Collapse
  - Behaves like a single queue
- Unused service under SSC generalizes:

$$q(k+1)u(k) = 0$$

$$\left(\sum_i q_i(k+1)\right)\left(\sum_i u_i\right) \approx 0$$

$$(e^{\theta\epsilon q(k+1)} - 1)(e^{-\theta\epsilon u(k)} - 1) = 0$$

$$E\left[(e^{\theta\epsilon \sum_i q_i(k+1)} - 1)(e^{-\theta\epsilon \sum_i u_i(k)} - 1)\right] \text{ is } o(\epsilon^2)$$

# CONCLUSIONS

- **Heavy-Traffic Analysis is easy!**

- Set drift of a test function to zero
- CRP: Exponential test function – Transform techniques

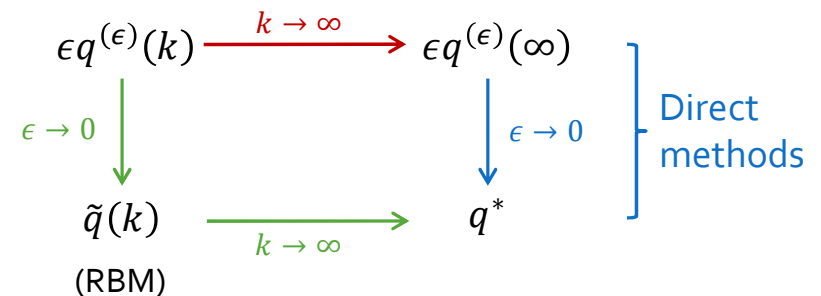
- **More than just heavy-traffic**

- Non-asymptotic results and Rates of Convergence
- Benefits of a 'Direct Method'

- **It's all about the unused service**

- **Future work**

- Extend MGF method to queueing systems that do not satisfy the CRP condition



# TRANSFORM METHODS FOR HEAVY TRAFFIC ANALYSIS

---

Daniela Hurtado Lange, Siva Theja Maguluri

Industrial and Systems Engineering

Georgia Institute of Technology

