Why are reflected random walks *so hard to simulate* ?

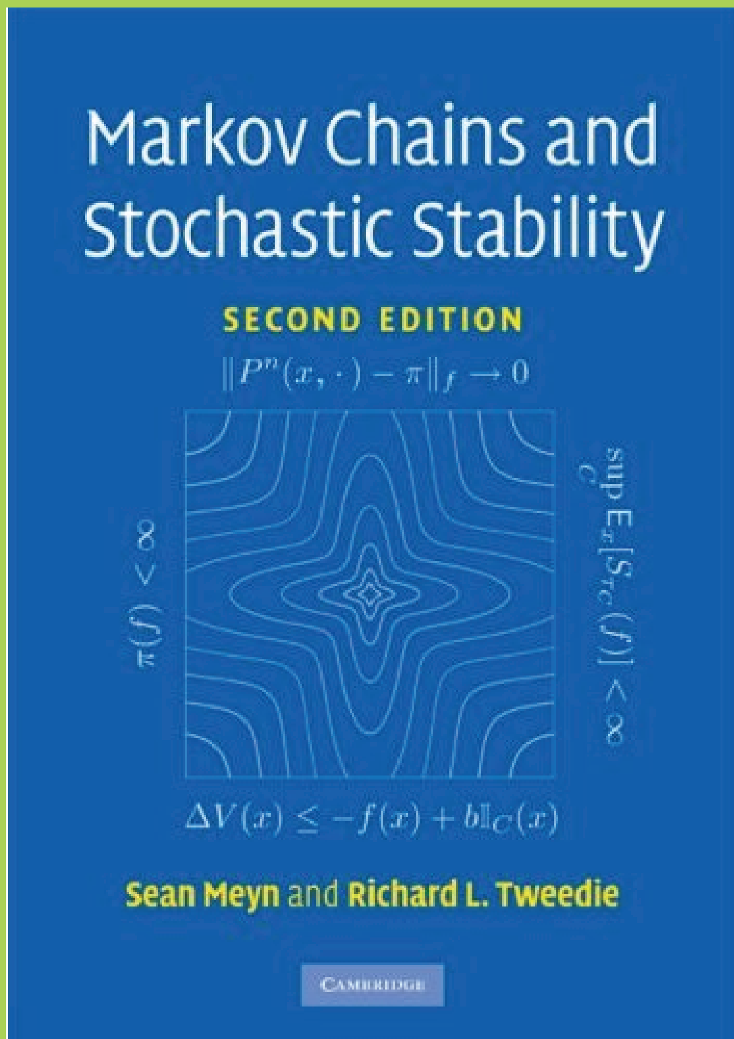
Sean Meyn

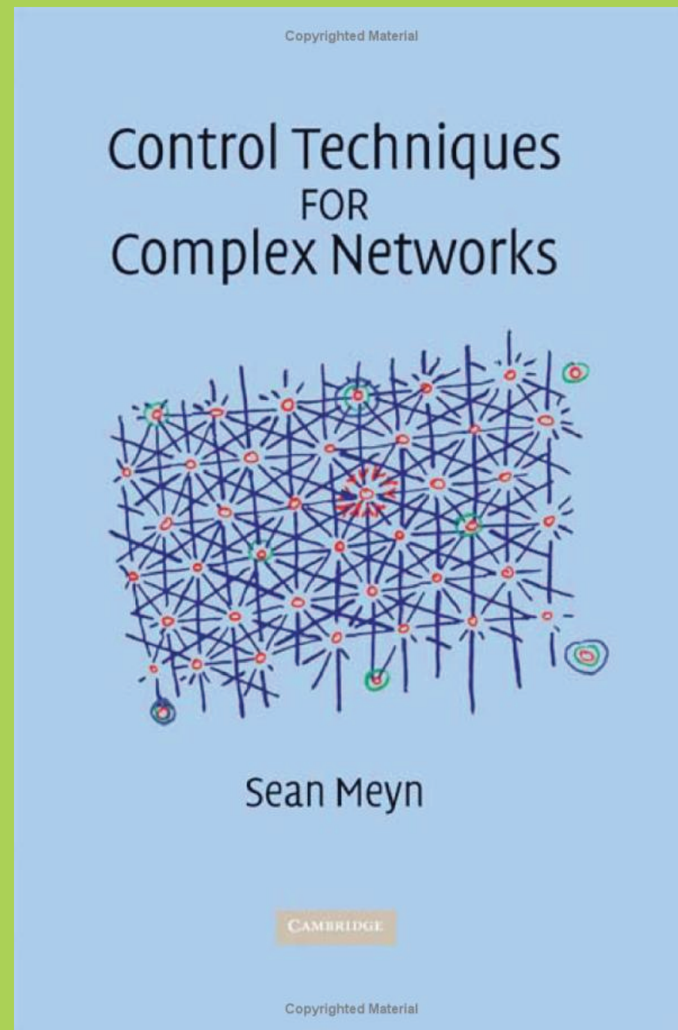
Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois

Joint work with Ken Duffy - Hamilton Institute, National University of Ireland

NSF support: ECS-0523620

References

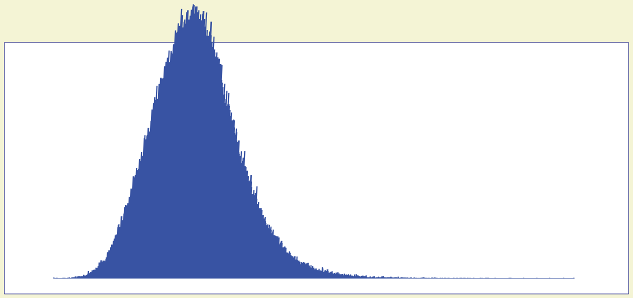




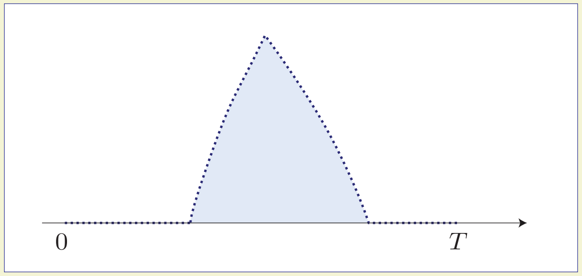
(Skip this advertisement)

Outline

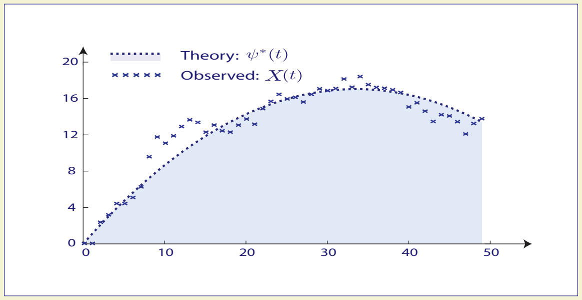
Motivation



Reflected Random Walks



Why So Lopsided?



Most Likely Paths

Summary and Conclusions

MM1 queue - the nicest of Markov chains

For the stable queue, with load strictly less than one, it is

- Reversible
- Skip-free
- Monotone
- Marginal distribution geometric
- Geometrically ergodic

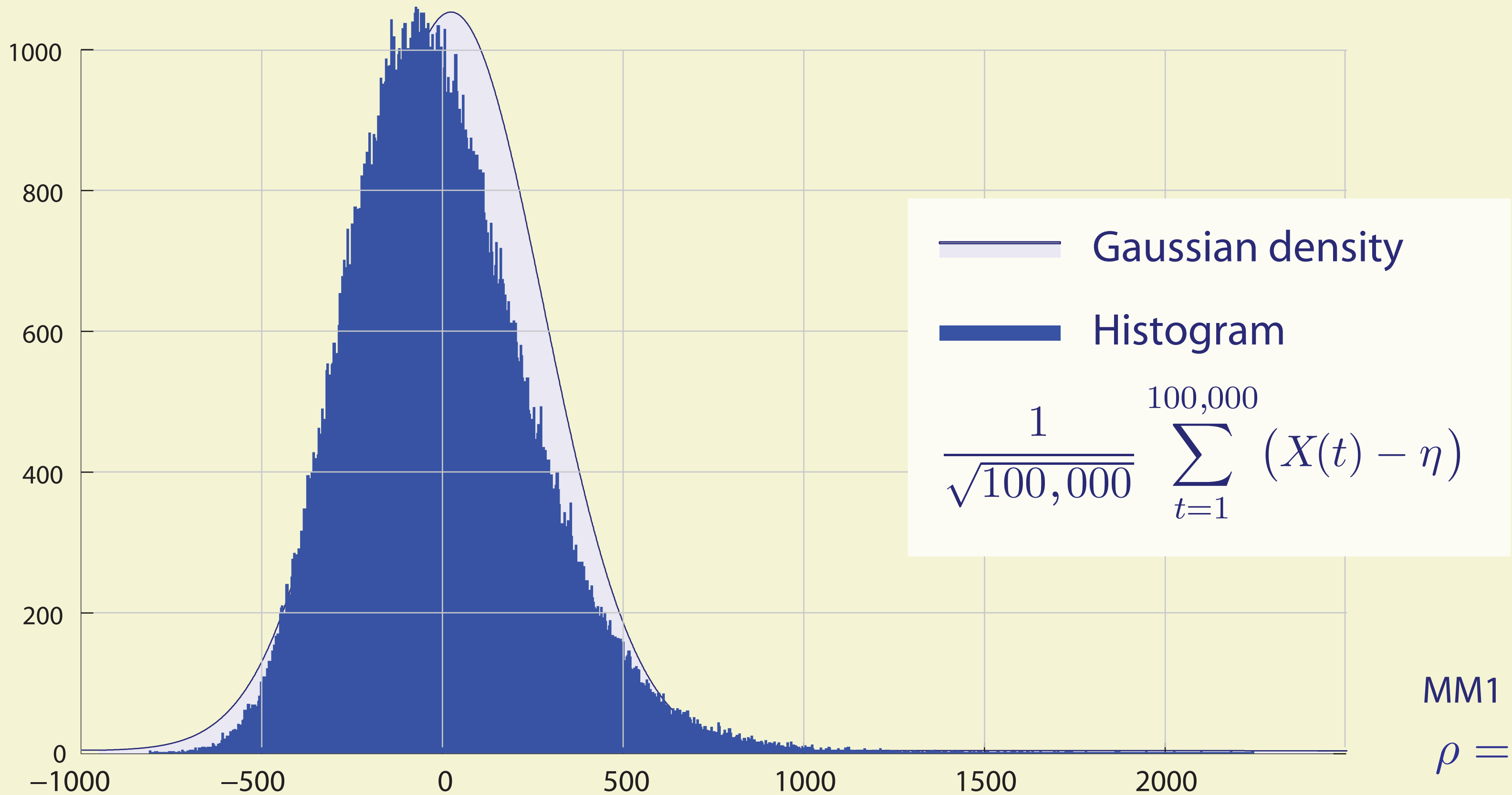
MM1 queue - the nicest of Markov chains

For the stable queue, with load strictly less than one, it is

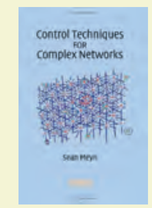
- Reversible
- Skip-free
- Monotone
- Marginal distribution geometric
- Geometrically ergodic

Why then, can't I simulate my queue?

MM1 queue - the nicest of Markov chains



Ch 11 CTCN

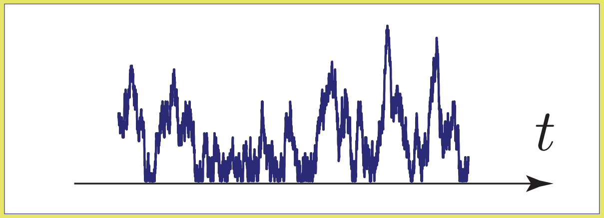


MM1 queue

$\rho = 0.9$

Why then, can't I simulate my queue?

II
Reflected Random Walk



Reflected Random Walk

A reflected random walk

$$X(n+1) = \max(0, X(n) + \Delta(n+1)),$$

where $X(0) = x_0 \in \mathbb{R}_+$ is the initial condition,
and the increments $\boldsymbol{\Delta}$ are i.i.d.

Reflected Random Walk

A reflected random walk

$$X(n+1) = \max(0, X(n) + \Delta(n+1)),$$

where $X(0) = x_0 \in \mathbb{R}_+$ is the initial condition,
and the increments $\boldsymbol{\Delta}$ are i.i.d.

Basic stability assumption: $\delta := \mathsf{E}[\Delta(n)] < 0$

and finite second moment

MM1 queue - the nicest of Markov chains

Reflected random walk with increments,

$$\Delta(n) = \begin{cases} 1 & \text{with prob. } \alpha \\ -1 & \text{with prob. } \mu \end{cases}$$

$$\rho = \alpha/\mu < 1$$

Load condition

Reflected Random Walk

A reflected random walk

$$X(n+1) = \max(0, X(n) + \Delta(n+1)),$$

where $X(0) = x_0 \in \mathbb{R}_+$ is the initial condition,
and the increments Δ are i.i.d.

Basic stability assumption: $\delta := \mathsf{E}[\Delta(n)] < 0$

Basic question: Can I compute sample path averages to estimate the steady-state mean?

$$\eta := \lim_{n\to\infty} \mathsf{E}[X(n)] \qquad \eta(n) = \frac{1}{n}\sum_{t=0}^{n} X(t)$$

Simulating the RRW: Asymptotic Variance

$$\eta(n) = \frac{1}{n}\sum_{t=0}^{n} X(t)$$

The CLT requires a third moment for the increments

$$\mathsf{E}[\Delta(0)] < 0 \text{ and } \mathsf{E}[\Delta(0)^3] < \infty$$

Asymptotic variance = $O\Big(\frac{1}{(1-\rho)^4}\Big)$

Whitt 1989
Asmussen 1992

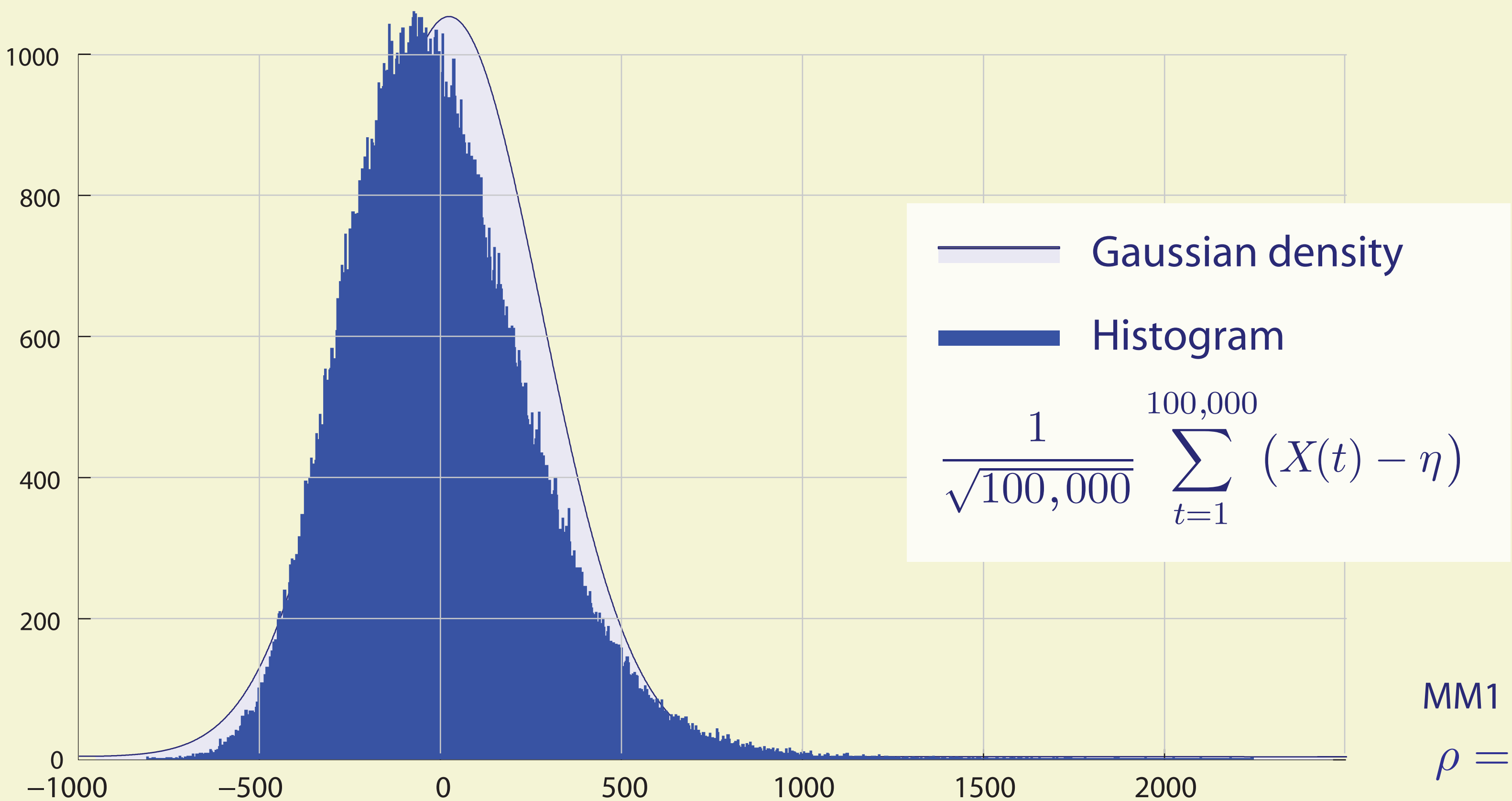
Simulating the RRW: Asymptotic Variance

$$\eta(n) = \frac{1}{n}\sum_{t=0}^{n} X(t)$$

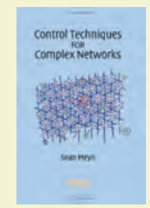
The CLT requires a third moment for the increments

$$\mathsf{E}[\Delta(0)] < 0 \text{ and } \mathsf{E}[\Delta(0)^3] < \infty$$

Asymptotic variance $= O\Big(\frac{1}{(1-\rho)^4}\Big)$



Ch 11 CTCN



MM1 queue

$\rho = 0.9$

$$\eta(n) = \frac{1}{n}\sum_{t=0}^{n} X(t)$$

Simulating the RRW: Lopsided Statistics

Assume only negative drift, and finite second moment:

$$\mathsf{E}[\Delta(0)] < 0 \text{ and } \mathsf{E}[\Delta(0)^2] < \infty$$

Lower LDP asymptotics: For each $r < \eta$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\eta(n) \leq r\} = -I(r) < 0$$

M 2006

$$\eta(n) = \frac{1}{n}\sum_{t=0}^{n} X(t)$$

Simulating the RRW: Lopsided Statistics

Assume only negative drift, and finite second moment:

$$\mathsf{E}[\Delta(0)] < 0 \text{ and } \mathsf{E}[\Delta(0)^2] < \infty$$

Lower LDP asymptotics: For each $r < \eta$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\eta(n) \le r\} = -I(r) < 0$$

M 2006

Upper LDP asymptotics are *null*: For each $r \ge \eta$,

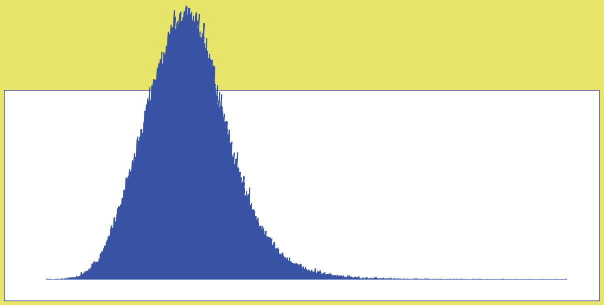
$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\eta(n) \ge r\} = 0$$

M 2006
CTCN

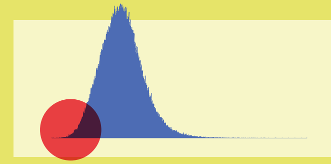
even for the MM1 *queue*!

III

Why So Lopsided?



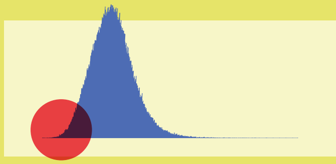
Lower LDP



Construct the family of *twisted* transition laws

$$\check{P}(x, dy) = e^{\theta x - \Lambda(\theta) + \check{F}(x)} P(x, dy) e^{\check{F}(y)} \qquad \theta < 0$$

Lower LDP



Construct the family of *twisted* transition laws

$$\check{P}(x,dy) = e^{\theta x - \Lambda(\theta) + \check{F}(x)} P(x,dy) e^{\check{F}(y)} \qquad \theta < 0$$

Geometric ergodicity of the twisted chain implies multiplicative ergodic theorems, and thence Bahadur-Rao LDP asymptotics

Kontoyiannis & M 2003, 2005
M 2006

Lower LDP

Construct the family of *twisted* transition laws

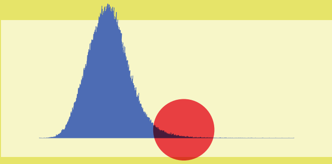
$$\check{P}(x, dy) = e^{\theta x - \Lambda(\theta) + \check{F}(x)} P(x, dy) e^{\check{F}(y)} \qquad \theta < 0$$

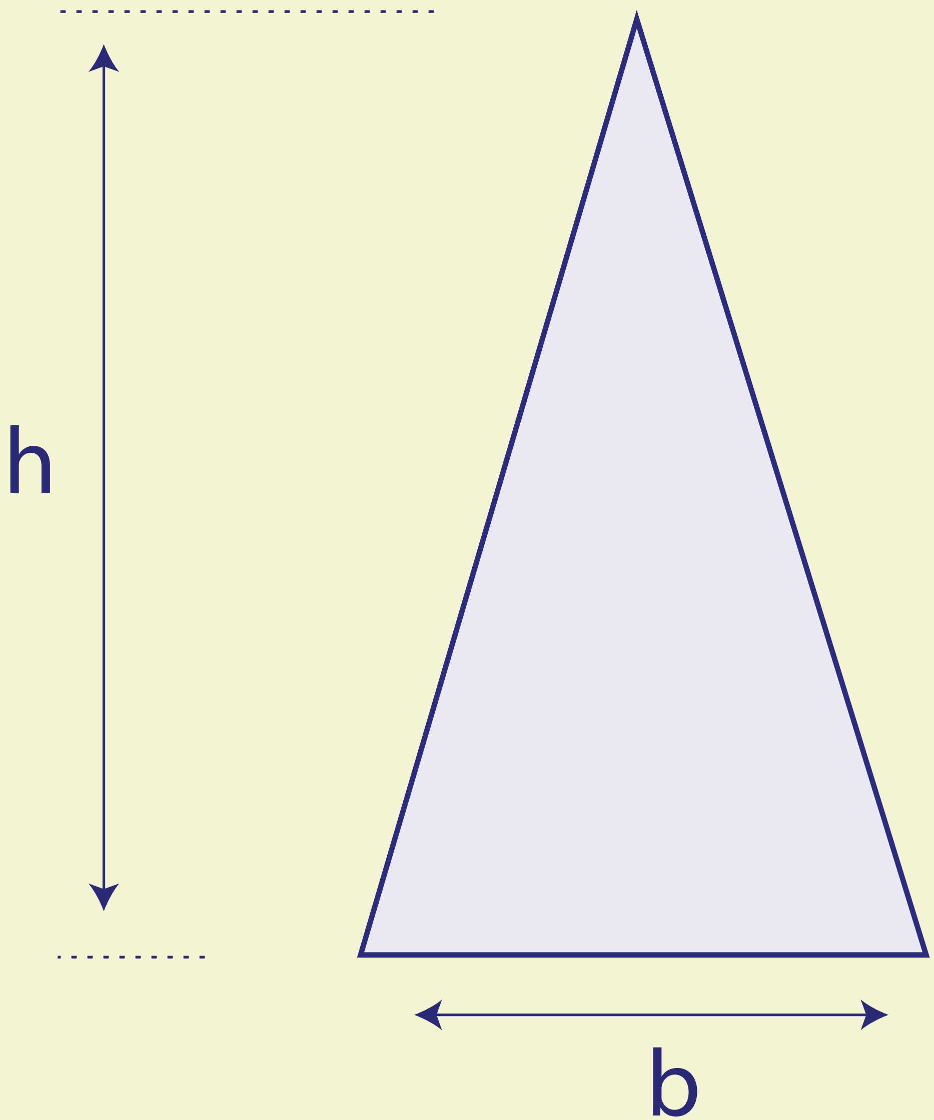
Geometric ergodicity of the twisted chain implies multiplicative ergodic theorems, and thence Bahadur-Rao LDP asymptotics

Kontoyiannis & M 2003, 2005
M 2006

This is only possible when θ *is negative*

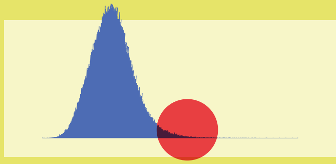
Null Upper LDP: What is the area of a triangle?

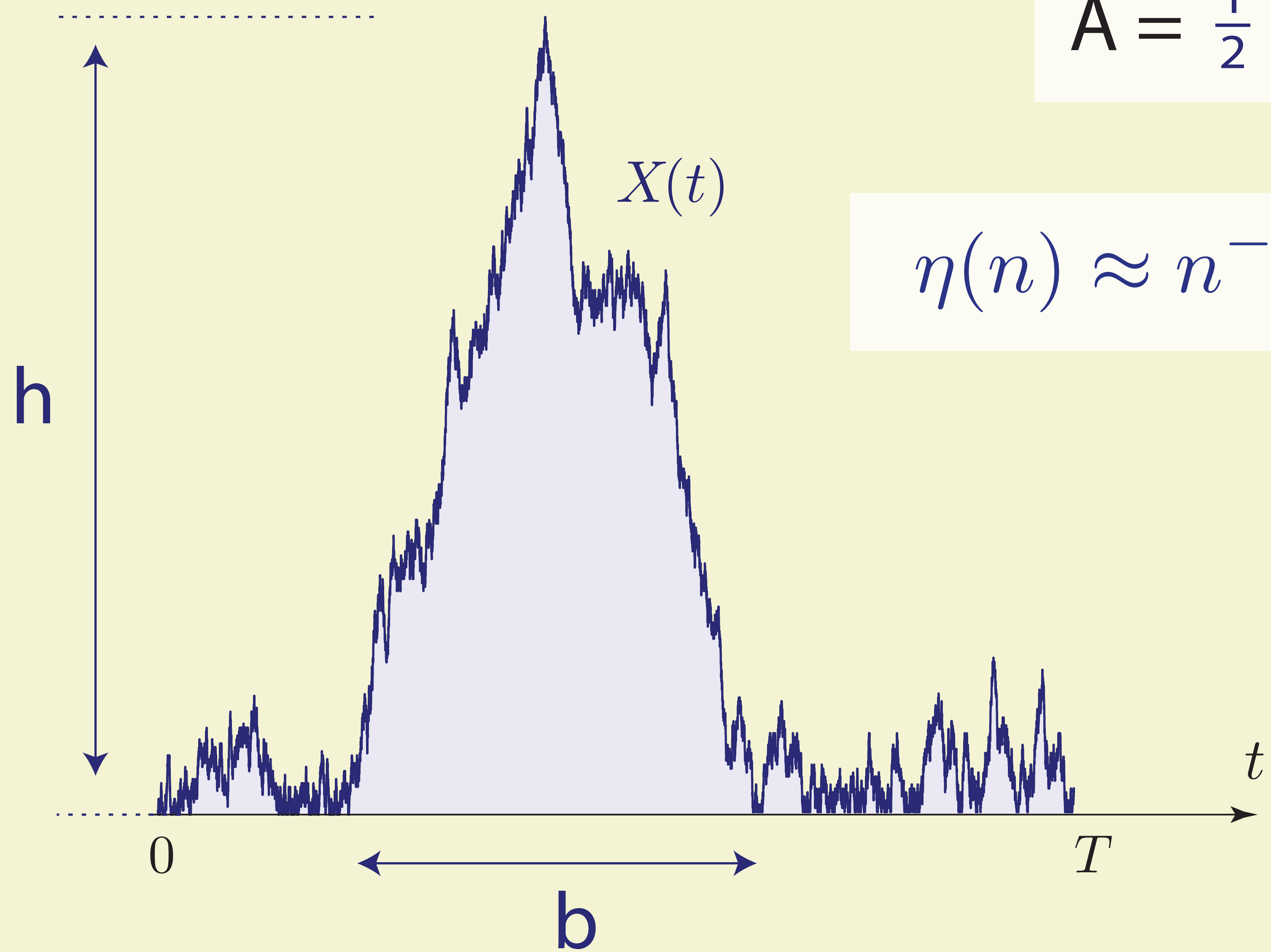




$$A = \frac{1}{2} bh$$

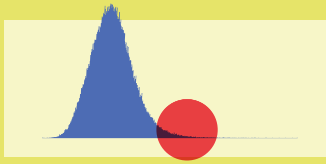
Null Upper LDP: *What is the area of a triangle?*





$$A = \frac{1}{2} bh$$

$$\eta(n) \approx n^{-1} A$$



Null Upper LDP: *Area of a Triangle?*

Consider the scaling $T = n$

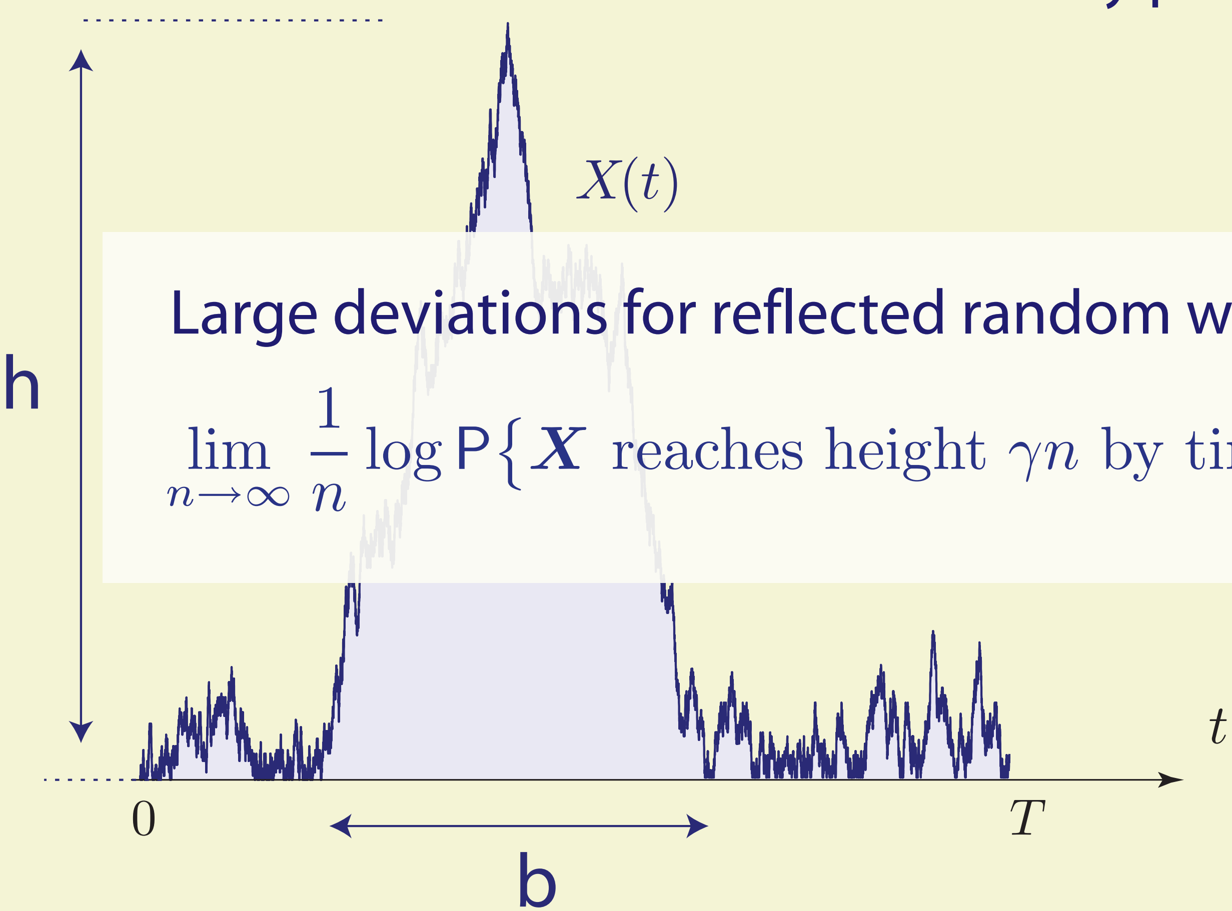
$$h = \gamma n$$

$$\eta(n) \approx n^{-1} \mathrm{A}$$

Base obtained from most likely path to this height:

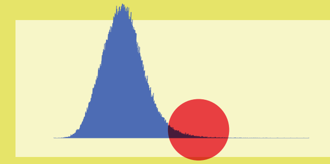
$$b = \beta^* n$$

e.g., Ganesh, O'Connell, and Wischik, 2004



Large deviations for reflected random walk:

$$\lim_{n \to \infty} \frac{1}{n} \log \mathrm{P}\{\boldsymbol{X} \text{ reaches height } \gamma n \text{ by time } n\} = -I(\gamma)$$



Null Upper LDP: *Area of a Triangle*

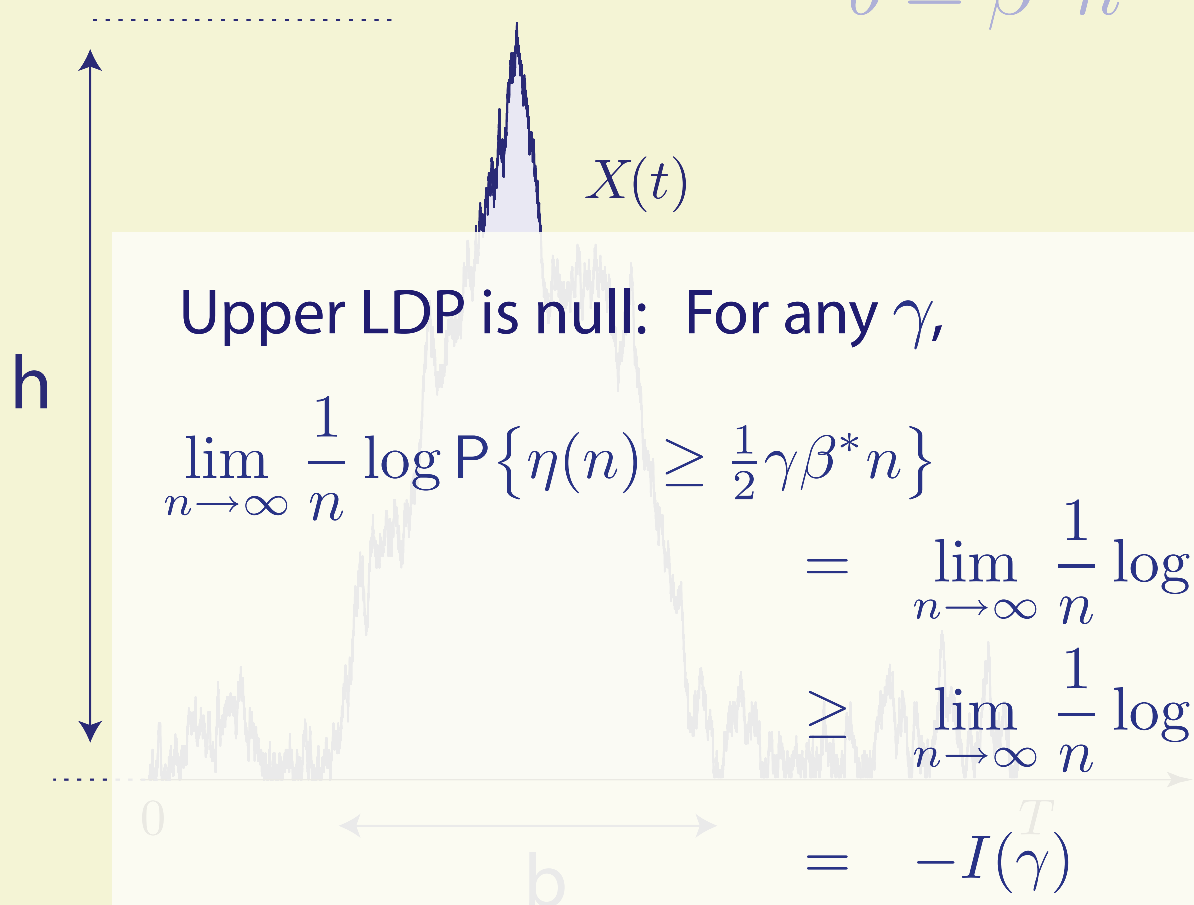
Consider the scaling $T = n$

$$h = \gamma n$$

$$b = \beta^* n$$

$$\eta(n) \approx n^{-1} \mathrm{A}$$

$$A_n = \tfrac{1}{2}\gamma\beta^* n^2$$

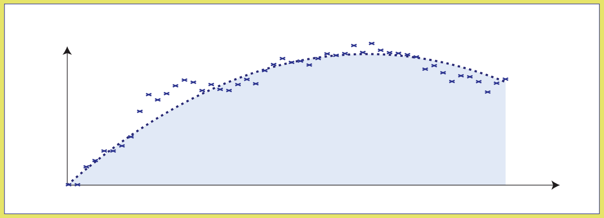


Upper LDP is null: For any γ,

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\eta(n) \ge \tfrac{1}{2}\gamma\beta^* n\} &= \lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\eta(n) \ge n^{-1} A_n\} \\
&\ge \lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\boldsymbol{X} \text{ reaches height } \gamma n\} \\
&= -I(\gamma)
\end{aligned}$$

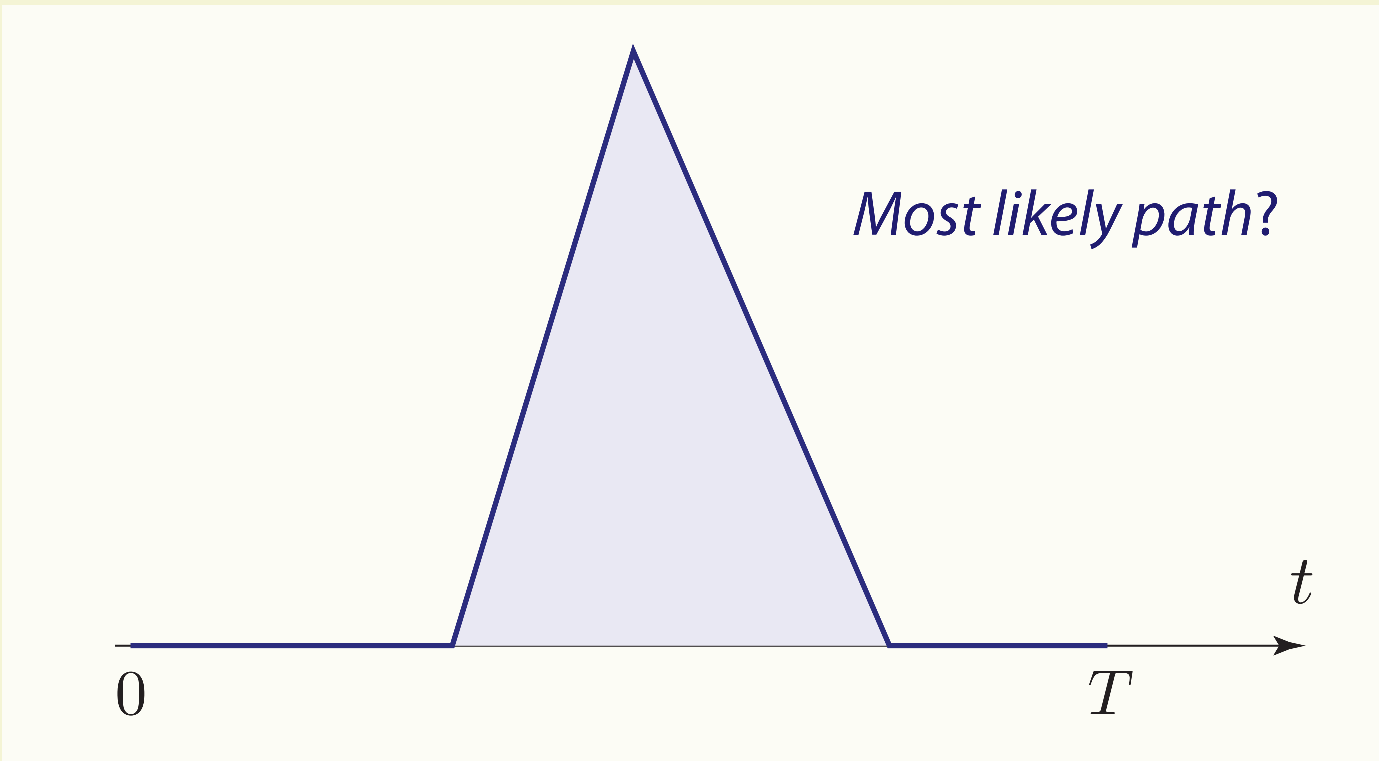
M 2006
CTCN 2008
(see also Borovkov, et. al. 2003)

IV
Most Likely Paths



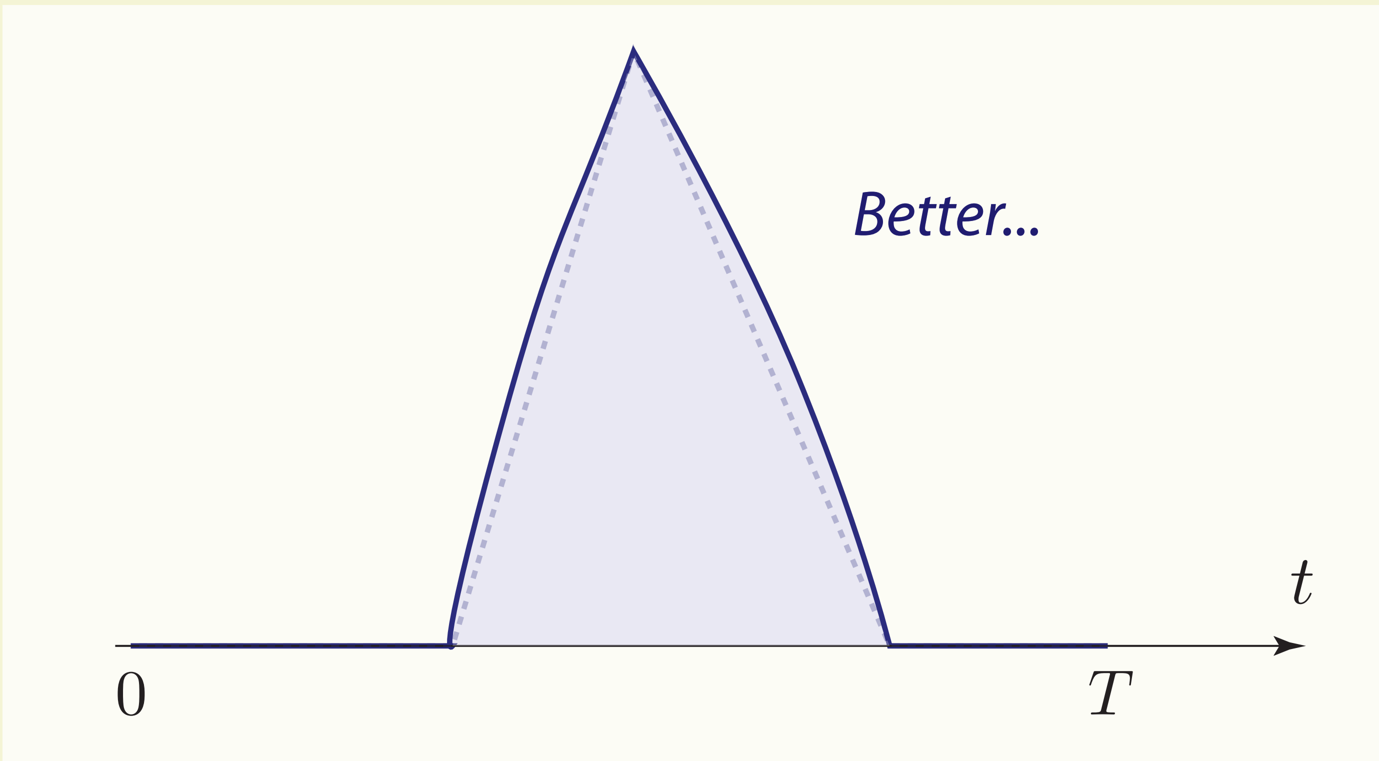
What Is The Most Likely Area?

Are triangular excursions optimal?



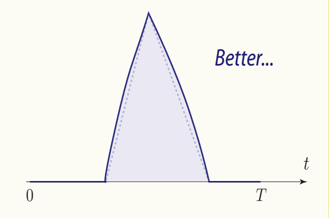
What Is The Most Likely Area?

Triangular excursions are *not* optimal:



A concave perturbation: greatly increased area at modest "cost"

Most Likely Paths



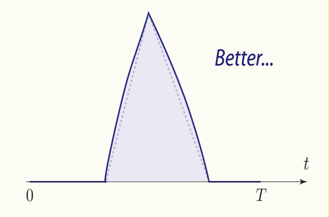
Scaled process

$$\psi^n(t) = n^{-1}X(nt)$$

LDP question translated to the scaled process

Duffy and M 2009, ...

Most Likely Paths



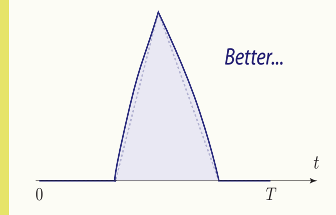
Scaled process

$$\psi^n(t) = n^{-1} X(nt)$$

LDP question translated to the scaled process

$$\eta(n) \approx An \Longleftrightarrow \int_0^1 \psi^n(t)\, dt \approx A$$

Most Likely Paths



Scaled process

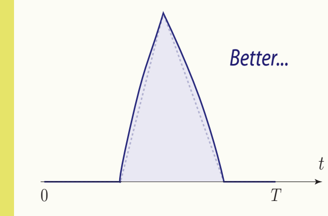
$$\psi^n(t) = n^{-1} X(nt)$$

LDP question translated to the scaled process

$$\eta(n) \approx An \Longleftrightarrow \int_0^1 \psi^n(t)\, dt \approx A$$

Basic assumption: The sample paths for the unconstrained random walk with increments Δ satisfy the LDP in $D[0,1]$ with good rate function I_X

Most Likely Paths



Scaled process

$$\psi^n(t) = n^{-1}X(nt)$$

LDP question translated to the scaled process

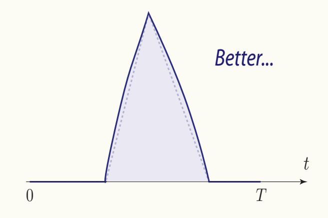
$$\eta(n) \approx An \iff \int_0^1 \psi^n(t)\,dt \approx A$$

Basic assumption: The sample paths for the unconstrained random walk with increments Δ satisfy the LDP in $D[0,1]$ with good rate function I_X

$$I_X(\psi) = \bar{\vartheta}^+ \psi(0+) + \int_0^1 I_\Delta(\dot{\psi}(t))\,dt$$

For *concave* ψ, with no downward jumps

Most Likely Path Via Dynamic Programming

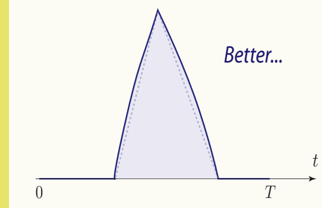


Dynamic programming formulation

$$\min \qquad \bar{\vartheta}^{+}\psi(0+) + \int_0^1 I_\Delta(\dot{\psi}(t))\,dt$$

$$\text{s.t.} \qquad \int_0^1 \psi(t)\,dt = A$$

Most Likely Path Via Dynamic Programming



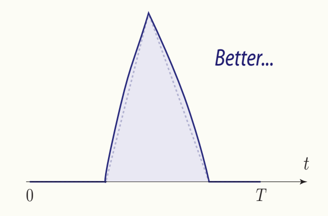
Dynamic programming formulation

$$\begin{aligned} \min \quad & \bar{\vartheta}^{+} \psi(0+) + \int_0^1 I_{\Delta}(\dot{\psi}(t))\, dt \\ \text{s.t.} \quad & \int_0^1 \psi(t)\, dt = A \end{aligned}$$

Solution: Integration by parts + Lagrangian relaxation:

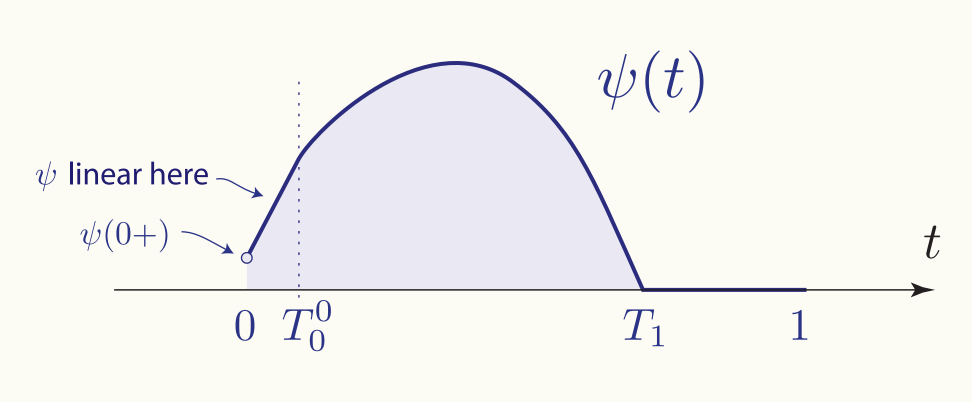
$$\min \ \bar{\vartheta}^{+} \psi(0+) + \int_0^1 I_{\Delta}(\dot{\psi}(t))\, dt + \lambda\Big(\psi(1) - A - \int_0^1 t\dot{\psi}(t)\, dt\Big)$$

Most Likely Path Via Dynamic Programming



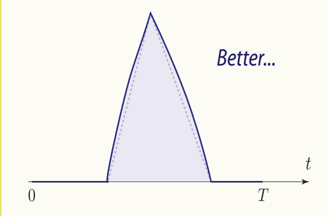
$$\min\ \bar{\vartheta}^{+}\,\psi(0+) + \int_0^1 I_\Delta(\dot{\psi}(t))\,dt + \lambda\Big(\psi(1) - A - \int_0^1 t\dot{\psi}(t)\,dt\Big)$$

Variational arguments where ψ is strictly concave:



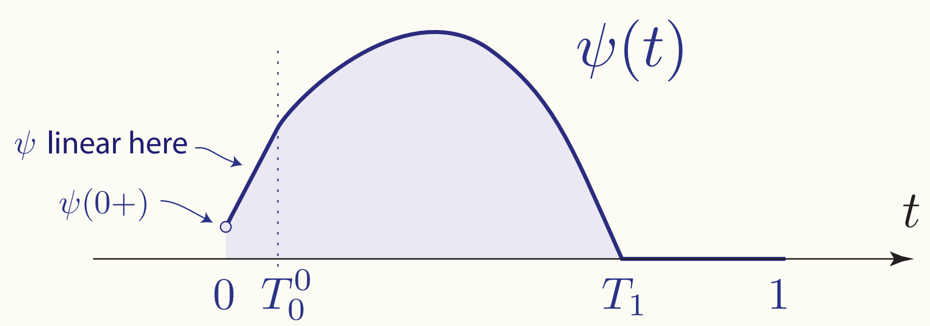
Strictly concave on (T_0^0, T_1)

Most Likely Path Via Dynamic Programming



$$\min\ \bar{\vartheta}^{+}\,\psi(0+) + \int_0^1 I_\Delta(\dot\psi(t))\,dt + \lambda\Big(\psi(1) - A - \int_0^1 t\dot\psi(t)\,dt\Big)$$

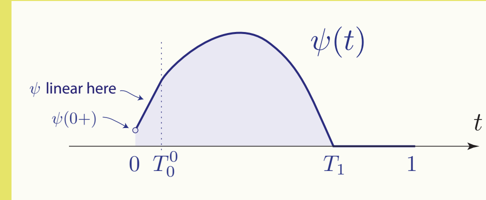
Variational arguments where ψ is strictly concave:

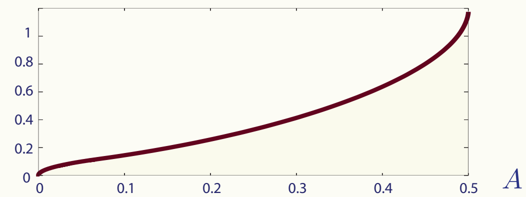
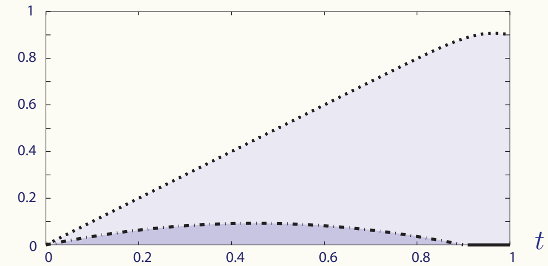
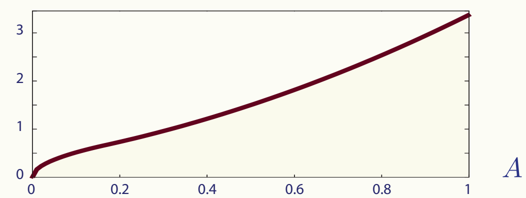
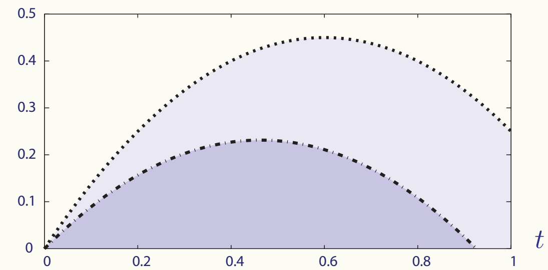
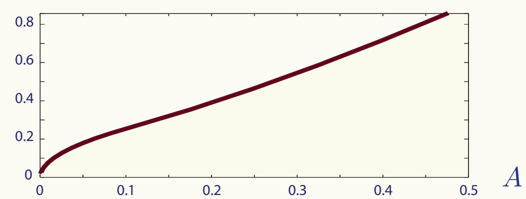
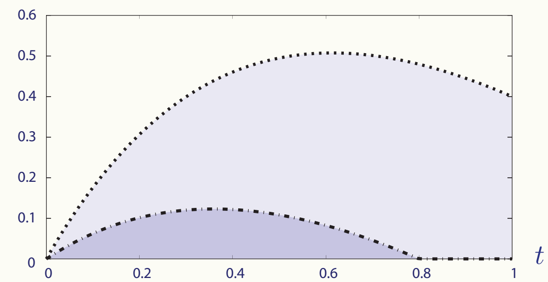
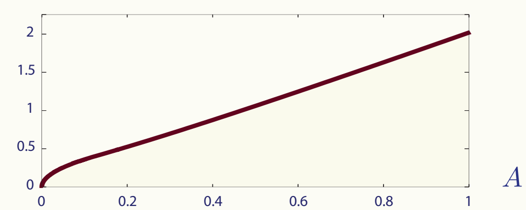
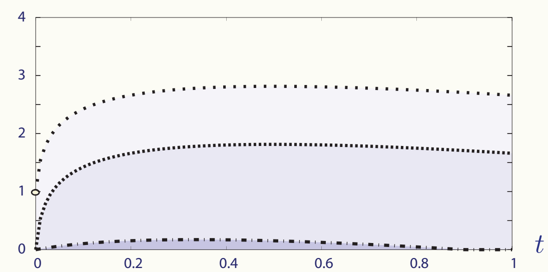


Strictly concave on (T_0^0, T_1)

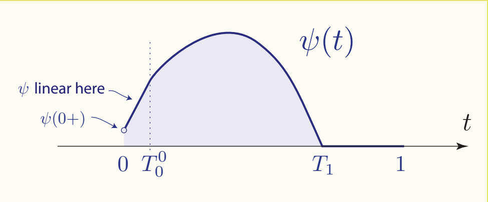
$$\nabla I(\dot\psi(t)) = b - \lambda^* t \quad \text{for } a.e.\ t \in (T_0^0, T_1)$$

Most Likely Paths - Examples

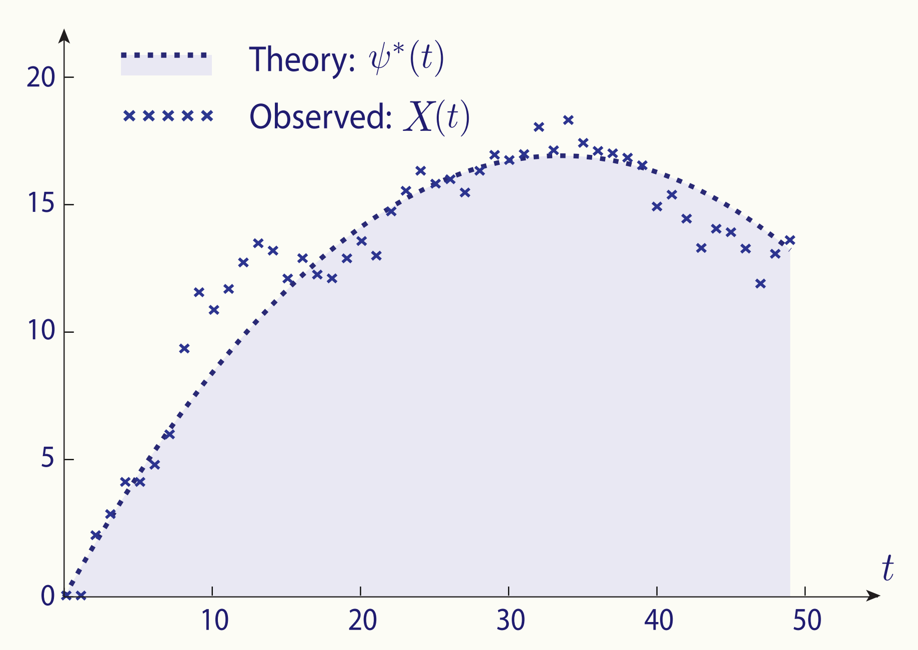


Increment	Exponent: $I^{+}\psi^{*}(0+) + \int_0^1 I_{\Delta}(\dot{\psi}^{*}(t))\,dt$	Optimal paths $\psi^{*}(t)$ for various A
± 1 (MM1 Queue)		
Gaussian		
Poisson		
Geometric		

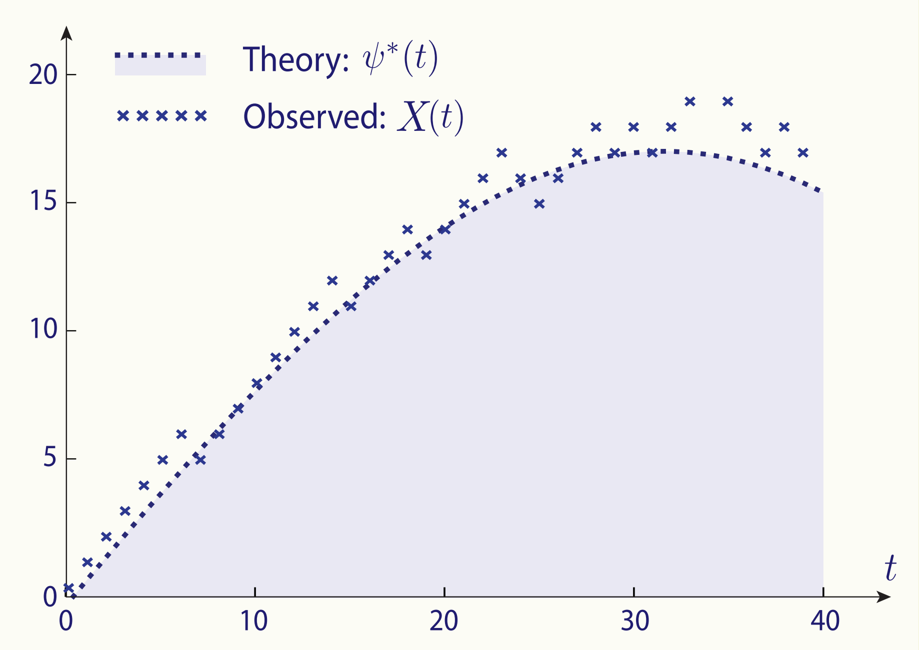
Most Likely Paths - Examples



Selected Sample Paths:



RRW: Gaussian Increments



RRW: MM1 queue

The observed path has the largest simulated mean out of 10^8 sample paths

Conclusions

Summary

It is widely known that simulation variance
is high in "heavy traffic" for queueing models

Large deviation asymptotics are exotic,
regardless of load

Sample-path behavior is identified for RRW when
the sample mean is large.
*This behavior is very different than previously seen
in "buffer overflow" analysis of queues*

Summary

It is widely known that simulation variance
is high in “heavy traffic” for queueing models

Large deviation asymptotics are exotic,
regardless of load

Sample-path behavior is identified for RRW when
the sample mean is large.
*This behavior is very different than previously seen
in “buffer overflow” analysis of queues*

What about other Markov models?

MM1 queue - the nicest of Markov chains

$$\Delta(n) = \begin{cases} 1 & \text{with prob. } \alpha \\ -1 & \text{with prob. } \mu \end{cases}$$

What about other Markov models?

- Reversible
- Skip-free

- Monotone
- Marginal distribution geometric
- Geometrically ergodic

MM1 queue - the nicest of Markov chains

$$\Delta(n) = \begin{cases} 1 & \text{with prob. } \alpha \\ -1 & \text{with prob. } \mu \end{cases}$$

What about other Markov models?

- Reversible
- Skip-free
- Monotone
- Marginal distribution geometric
- Geometrically ergodic

The skip-free property makes
the fluid model analysis possible. Similar behavior in

- Fixed-gain SA
- Some MCMC algorithms

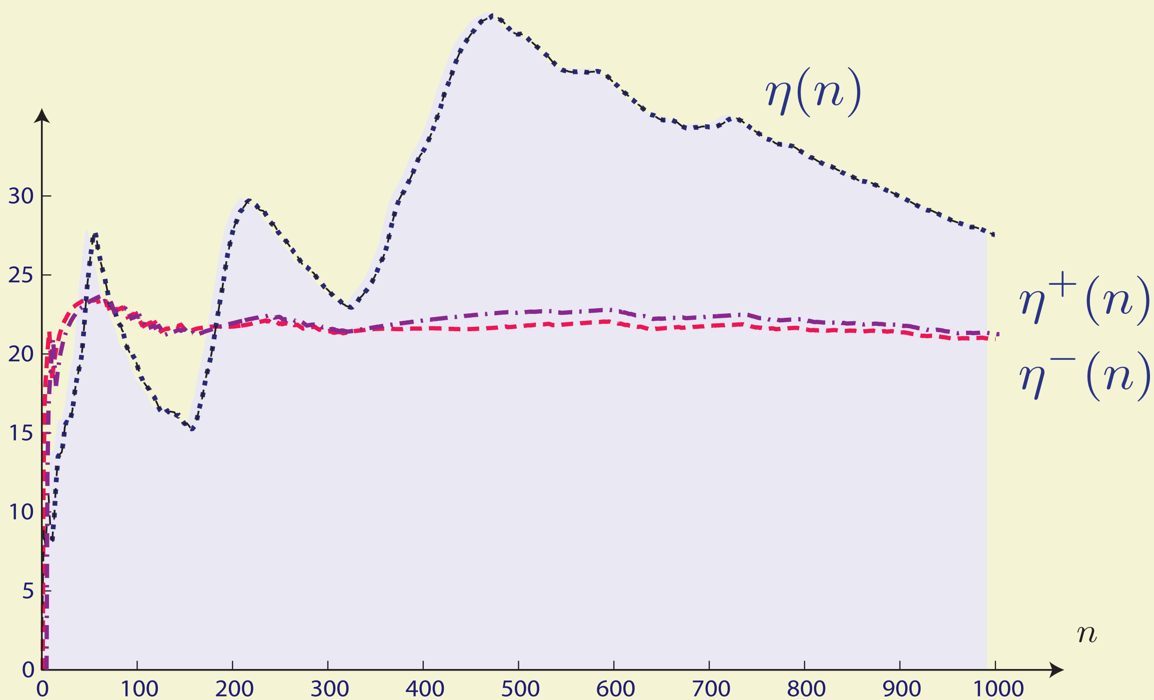
What Next?

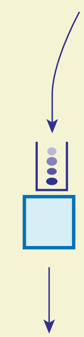
- *Heavy-tailed* and/or *long-memory* settings?

What Next?

- *Heavy-tailed* and/or *long-memory* settings?

- Variance reduction, such as control variates





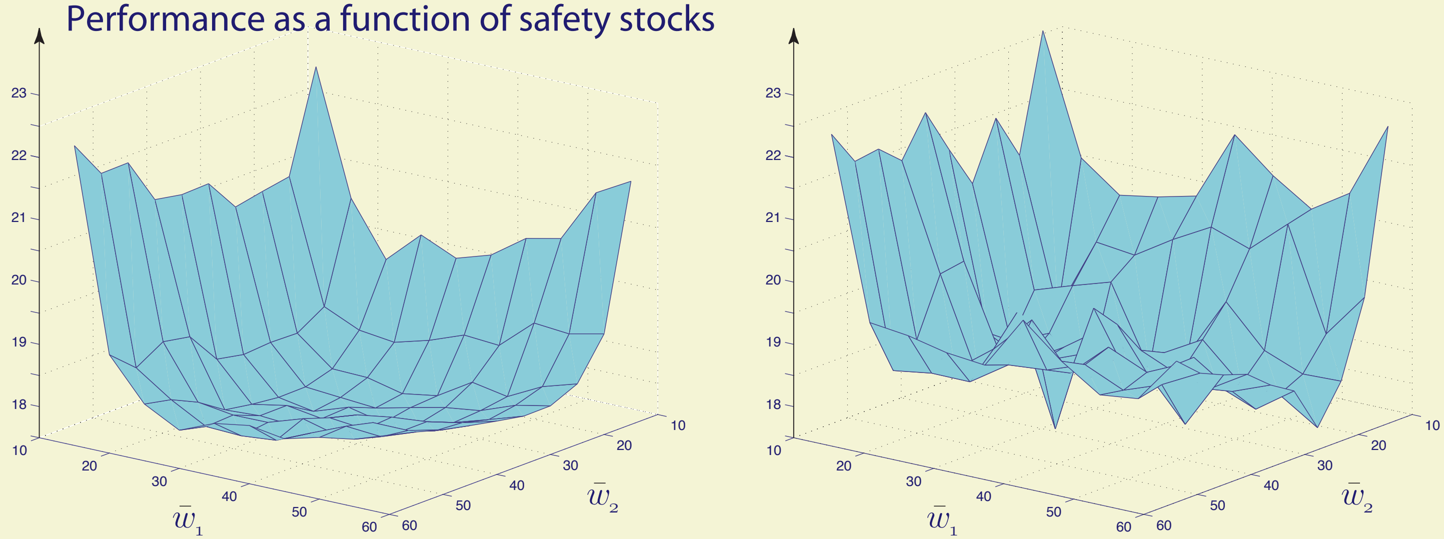
M 2006
CTCN

What Next?

- *Heavy-tailed* and/or *long-memory* settings?

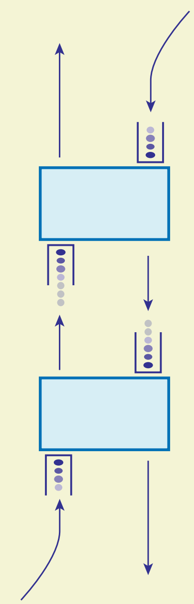
- Variance reduction, such as control variates

Performance as a function of safety stocks



(i) Simulation using smoothed estimator

(ii) Simulation using standard estimator



Smoothed estimator using fluid value function in CV: $g = h - Ph = -\mathcal{D}h, \quad h = \widehat{J}$

Henderson, M., and Tadi´c 2003
CTCN

What Next?

- *Heavy-tailed* and/or *long-memory* settings?
- Variance reduction, such as control variates
- Original question? Conjecture,

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \log \mathsf{P}\{\eta(n) \geq r\} = -J_\eta(r) < 0, \qquad r > \eta$$

for some range of r

References

LDPs for RRW

[1] S. P. Meyn. Large deviation asymptotics and control variates for simulating large functions. *Ann. Appl. Probab.*, 16(1):310–339, 2006.

[2] S. P. Meyn. *Control Techniques for Complex Networks.* Cambridge University Press, Cambridge, 2007.

[3] K. R. Duffy and S. P. Meyn. Most likely paths to error when estimating the mean of a reflected random walk. http://arxiv.org/abs/0906.4514, June 2009.

Exact LDPs

[1] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13:304–362, 2003.

[2] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10(3):61–123 (electronic), 2005.

Simulation

[1] S. G. Henderson, S. P. Meyn, and V. B. Tadić. Performance evaluation and policy selection in multiclass networks. *Discrete Event Dynamic Systems: Theory and Applications*, 13(1-2):149–189, 2003. Special issue on learning, optimization and decision making (invited).

[2] S. P. Meyn. *Control Techniques for Complex Networks.* Cambridge University Press, Cambridge, 2007.

[1] G. Fort, S. Meyn, E. Moulines, and P. Priouret. ODE methods for skip-free Markov chain stability with applications to MCMC. *Ann. Appl. Probab.*, 18(2):664–707, 2008.

[2] V. S. Borkar and S. P. Meyn. The ODE method for convergence of stochastic approximation and reinforcement learning. *SIAM J. Control Optim.*, 38(2):447–469, 2000.