On the design and analysis of branching schemes with killing for rare event Monte Carlo estimation

Paul Dupuis

(with Thomas Dean, Oxford University)
Division of Applied Mathematics
Brown University

Rubinstein Conference at Sandbjerg Estate

July 2008

Outline

1. Splitting-type schemes

Outline

1. Splitting-type schemes
2. RESTART (REpetitive Simulation Trials After Reaching Thresholds) and DPR (Direct Probability Redistribution)

Outline

1. Splitting-type schemes
2. RESTART (REpetitive Simulation Trials After Reaching Thresholds) and DPR (Direct Probability Redistribution)
3. Representations and bounds for the second moment

Outline

1. Splitting-type schemes
2. RESTART (REpetitive Simulation Trials After Reaching Thresholds) and DPR (Direct Probability Redistribution)
3. Representations and bounds for the second moment
4. Statement of main results

Outline

1. Splitting-type schemes
2. RESTART (REpetitive Simulation Trials After Reaching Thresholds) and DPR (Direct Probability Redistribution)
3. Representations and bounds for the second moment
4. Statement of main results
5. The F-V quasipotential and subsolutions

Outline

1. Splitting-type schemes
2. RESTART (REpetitive Simulation Trials After Reaching Thresholds) and DPR (Direct Probability Redistribution)
3. Representations and bounds for the second moment
4. Statement of main results
5. The F-V quasipotential and subsolutions
6. Remarks

Problem of Interest and LD Scaling

We consider a discrete time process $\{X_i^n\}$, not necessarily Markov.

Problem of Interest and LD Scaling

We consider a discrete time process $\{X_i^n\}$, not necessarily Markov. Let

$$X^n(i/n) = X_i^n, \quad \text{piecewise linear interpolation for } t \neq i/n.$$

Problem of Interest and LD Scaling

We consider a discrete time process $\{X_i^n\}$, not necessarily Markov. Let

$$X^n(i/n) = X_i^n, \quad \text{piecewise linear interpolation for } t \neq i/n.$$

Assume $\{X^n(\cdot)\}$ satisfies a Large Deviation Principle with rate function

$$I_T(\phi) = \int_0^T L(\phi, \dot{\phi}) dt$$

if ϕ is AC and $I_T(\phi) = \infty$ else.

Problem of Interest and LD Scaling

We consider a discrete time process $\{X_i^n\}$, not necessarily Markov. Let

$$X^n(i/n) = X_i^n, \quad \text{piecewise linear interpolation for } t \neq i/n.$$

Assume $\{X^n(\cdot)\}$ satisfies a Large Deviation Principle with rate function

$$I_T(\phi) = \int_0^T L(\phi, \dot{\phi})dt$$

if ϕ is AC and $I_T(\phi) = \infty$ else. Heuristically, for $T < \infty$, given ϕ, small $\delta > 0$, $x_n \to x = \phi(0)$ and large n

$$P_{x_n}\left\{\sup_{0\leq t\leq T} \|X^n(t) - \phi(t)\| \leq \delta\right\} \approx e^{-nI_T(\phi)}.$$

Problem of Interest and LD Scaling

To estimate:

$$E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right], \text{ where } \tau^n \doteq \inf\left\{i : X_i^n \in M\right\}.$$

Problem of Interest and LD Scaling

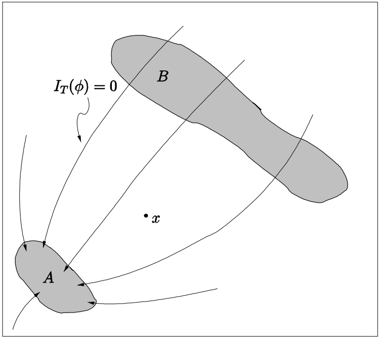
To estimate:

$$E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right], \text{ where } \tau^n \doteq \inf\left\{i : X_i^n \in M\right\}.$$

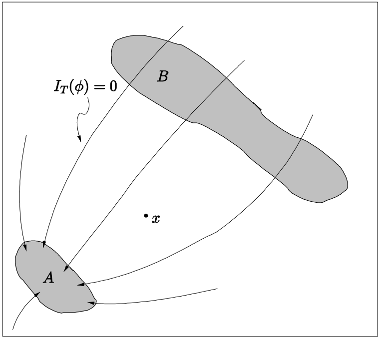
Example:

$M = A \cup B$, B rare, A typical, and $F(x) = 0$, $x \in B$, $F(x) = \infty$ otherwise.

Problem of Interest and LD Scaling



Problem of Interest and LD Scaling



Under conditions of regularity on F and bounds on τ^n:

$$
\begin{aligned}
-\frac{1}{n}\log E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right] &\rightarrow \inf\left\{I_T(\phi) + F(\phi(T)) : \phi(0) = x, T < \infty\right\} \\
&= \gamma.
\end{aligned}
$$

Some Estimation Generalities

1. General approach: construct iid random variables $\theta_1^n, \ldots, \theta_K^n$ with $E\theta_1^n = E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right]$ and use the unbiased estimator

$$\hat{q}_{n,K}(x_n) \doteq \frac{\theta_1^n + \cdots + \theta_K^n}{K}.$$

Some Estimation Generalities

1. General approach: construct iid random variables $\theta_1^n, \ldots, \theta_K^n$ with $E\theta_1^n = E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right]$ and use the unbiased estimator
$$\hat{q}_{n,K}(x_n) \doteq \frac{\theta_1^n + \cdots + \theta_K^n}{K}.$$
2. Performance determined by variance of θ_1^n, and since unbiased by $E\left(\theta_1^n\right)^2$.

Some Estimation Generalities

1. General approach: construct iid random variables $\theta_1^n, \ldots, \theta_K^n$ with $E\theta_1^n = E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right]$ and use the unbiased estimator

$$\hat{q}_{n,K}(x_n) \doteq \frac{\theta_1^n + \cdots + \theta_K^n}{K}.$$

2. Performance determined by variance of θ_1^n, and since unbiased by $E\left(\theta_1^n\right)^2$.
3. By Jensen's inequality

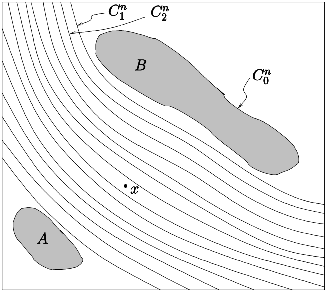
$$-\frac{1}{n}\log E\left(\theta_1^n\right)^2 \leq -\frac{2}{n}\log E\theta_1^n = -\frac{2}{n}\log E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right] \to 2\gamma.$$

Some Estimation Generalities

1. General approach: construct iid random variables $\theta_1^n, \ldots, \theta_K^n$ with $E\theta_1^n = E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right]$ and use the unbiased estimator
$$\hat{q}_{n,K}(x_n) \doteq \frac{\theta_1^n + \cdots + \theta_K^n}{K}.$$
2. Performance determined by variance of θ_1^n, and since unbiased by $E\left(\theta_1^n\right)^2$.
3. By Jensen's inequality
$$-\frac{1}{n}\log E\left(\theta_1^n\right)^2 \leq -\frac{2}{n}\log E\theta_1^n = -\frac{2}{n}\log E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-nF(X_i^n)}\right] \to 2\gamma.$$
4. An estimator is called *asymptotically efficient* if
$$\liminf_{n\to\infty} -\frac{1}{n}\log E\left(\theta_1^n\right)^2 \geq 2\gamma.$$

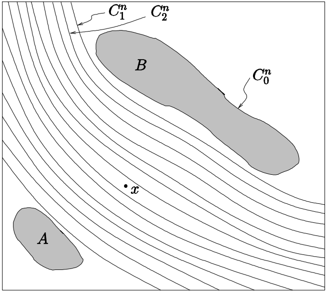
Splitting-Type Schemes

A certain number [proportional to n] of *splitting thresholds* C_r^n are defined which enhance migration, e.g., for hitting probabilities:



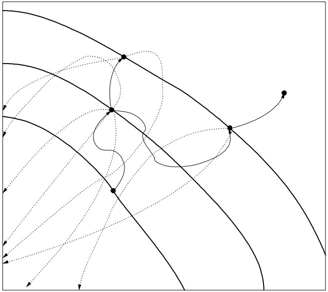
Splitting-Type Schemes

A certain number [proportional to n] of *splitting thresholds* C_r^n are defined which enhance migration, e.g., for hitting probabilities:

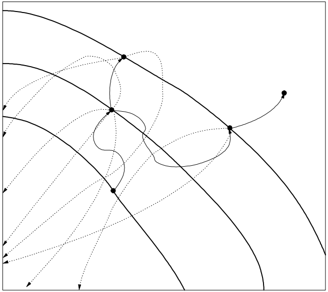


A single particle is started at x that follows the same law as X^n, but branches into a number of independent copies each time a new level is reached.

Splitting-type Schemes



Splitting-type Schemes



The number of new particles S can be random (though independent of past data), and a multiplicative weight w_i is assigned to the ith descendent, where

$$E\sum_{i=1}^{S} w_i = 1.$$

Splitting-Type Schemes

Evolution continues until every particle hits M. Let

$$\begin{aligned} N_x^n &= \text{number of particles generated} \\ X_i^n(j) &= \text{trajectory of } j\text{th particle,} \\ W_i^n(j) &= \text{product of weights assigned to } j \text{ along path up to time } i \\ \tau^n(j) &= \text{hitting time of } j\text{th trajectory} \end{aligned}$$

Splitting-Type Schemes

Evolution continues until every particle hits M. Let

$$\begin{aligned} N_x^n &= \text{number of particles generated} \\ X_i^n(j) &= \text{trajectory of } j\text{th particle,} \\ W_i^n(j) &= \text{product of weights assigned to } j \text{ along path up to time } i \\ \tau^n(j) &= \text{hitting time of } j\text{th trajectory} \end{aligned}$$

Then

$$\theta^n = \sum_{j=1}^{N_x^n} \sum_{i=0}^{\tau^n(j)} e^{-nF(X_i^n(j))} W_i^n(j).$$

Splitting-Type Schemes

Evolution continues until every particle hits M. Let

$$
\begin{aligned}
N_x^n &= \text{number of particles generated} \\
X_i^n(j) &= \text{trajectory of } j\text{th particle,} \\
W_i^n(j) &= \text{product of weights assigned to } j \text{ along path up to time } i \\
\tau^n(j) &= \text{hitting time of } j\text{th trajectory}
\end{aligned}
$$

Then

$$\theta^n = \sum_{j=1}^{N_x^n} \sum_{i=0}^{\tau^n(j)} e^{-nF(X_i^n(j))} W_i^n(j).$$

Problems

- If splitting is too infrequent, do not explore state space (standard Monte Carlo).

Splitting-Type Schemes

Evolution continues until every particle hits M. Let

$$\begin{aligned} N_x^n &= \text{number of particles generated} \\ X_i^n(j) &= \text{trajectory of } j\text{th particle,} \\ W_i^n(j) &= \text{product of weights assigned to } j \text{ along path up to time } i \\ \tau^n(j) &= \text{hitting time of } j\text{th trajectory} \end{aligned}$$

Then

$$\theta^n = \sum_{j=1}^{N_x^n} \sum_{i=0}^{\tau^n(j)} e^{-nF(X_i^n(j))} W_i^n(j).$$

Problems

- If splitting is too infrequent, do not explore state space (standard Monte Carlo).
- If too frequent, we have exponential growth in number of surviving particles.

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
    - Particles can jump at most one threshold ($j \to j + 1$) at each time step.

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.
  - Splitting occurs with *every* upcrossing, and particles are killed when they leave the threshold in which they were born.

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.
  - Splitting occurs with *every* upcrossing, and particles are killed when they leave the threshold in which they were born.
- DPR (Direct Probability Redistribution, due to Haraszti and Townsend). Same as RESTART but

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.
  - Splitting occurs with *every* upcrossing, and particles are killed when they leave the threshold in which they were born.
- DPR (Direct Probability Redistribution, due to Haraszti and Townsend). Same as RESTART but
  - Particles can jump multiple thresholds ($j \to k$).

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.
  - Splitting occurs with *every* upcrossing, and particles are killed when they leave the threshold in which they were born.
- DPR (Direct Probability Redistribution, due to Haraszti and Townsend). Same as RESTART but
  - Particles can jump multiple thresholds ($j \to k$).
  - Offspring are assigned a killing threshold from $\{j+1, \ldots, k\}$ according to multinomial distribution, chosen to achieve unbiasedness.

RESTART and DPR

An obvious inefficiency–continuing trajectories far from the places of interest. When killing trajectories, care needed to avoid bias.

- RESTART (REpetitive Simulation Trials After Reaching Thresholds, due to Villen-Altamirano and Villen-Altamirano).
  - Particles can jump at most one threshold ($j \to j+1$) at each time step.
  - When branching there is one parent and remainder are offspring.
  - Splitting occurs with *every* upcrossing, and particles are killed when they leave the threshold in which they were born.
- DPR (Direct Probability Redistribution, due to Haraszti and Townsend). Same as RESTART but
  - Particles can jump multiple thresholds ($j \to k$).
  - Offspring are assigned a killing threshold from $\{j+1, \ldots, k\}$ according to multinomial distribution, chosen to achieve unbiasedness.
- Same problems as with ordinary splitting, but analysis much more difficult due to dependence on threshold of birth.

Implementation Via Importance Functions

Given a continuous function U and $\Delta > 0$ define thresholds by

Implementation Via Importance Functions

Given a continuous function U and $\Delta > 0$ define thresholds by

$$V^n(y) = \Delta\left(\left\lfloor \frac{nU(x_n) - nU(y)}{\Delta} \right\rfloor \vee 0\right)$$

Implementation Via Importance Functions

Given a continuous function U and $\Delta > 0$ define thresholds by

$$V^n(y) = \Delta\left(\left\lfloor \frac{nU(x_n) - nU(y)}{\Delta} \right\rfloor \vee 0\right)$$

$$C_r^n = \{y : V^n(y) \geq r\Delta\}$$

Implementation Via Importance Functions

Given a continuous function U and $\Delta > 0$ define thresholds by

$$V^n(y) = \Delta \left(\left\lfloor \frac{nU(x_n) - nU(y)}{\Delta} \right\rfloor \vee 0 \right)$$

$$C_r^n = \{y : V^n(y) \geq r\Delta\}$$

and a mean increase in number of particles per threshold of

$$e^{n\Delta}.$$

Representations and Bounds for the Second Moment

Exact representations for variance of estimator possible but opaque. Some useful bounds in terms of *original process*. Derived via dynamic programming type arguments:

Representations and Bounds for the Second Moment

Exact representations for variance of estimator possible but opaque. Some useful bounds in terms of *original process*. Derived via dynamic programming type arguments:

$$E\left[(\theta^n)^2\right] \geq E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-n\left(U(x_n)-U(X_i^n)\right)\vee 0+o(n)} e^{-n2F(X_i^n)}\right],$$

Representations and Bounds for the Second Moment

Exact representations for variance of estimator possible but opaque. Some useful bounds in terms of *original process*. Derived via dynamic programming type arguments:

$$E\left[(\theta^n)^2\right] \geq E_{x_n}\left[\sum_{i=0}^{\tau^n} e^{-n\left(U(x_n)-U(X_i^n)\right)\vee 0+o(n)} e^{-n2F(X_i^n)}\right],$$

$$\begin{aligned} E\left[(\theta^n)^2\right] \quad &\leq \quad E_{x_n}\Bigg[\sum_{i=0}^{\tau^n} e^{-n\left(U(x_n)-U(X_i^n)\right)\vee 0+o(n)} \\ &\qquad \times \left[\sum_{j=i}^{\tau^{1,i,n}} e^{-nF(X_j^{1,i,n})}\right]\left[\sum_{j=i}^{\tau^{2,i,n}} e^{-nF(X_j^{2,i,n})}\right]\Bigg]\end{aligned}$$

where $X_j^{k,i,n}$ are (conditionally) independent copies of X_j^n that start at X_i^n at $j = i$.

Statement of Main Results

Let

$$\mathcal{J}(y,z) = \inf\left\{ I_T(\phi) : \phi(0) = y, \phi(T) = z, T < \infty \right\}.$$

Statement of Main Results

Let

$$\mathcal{J}(y,z) = \inf \left\{ I_T(\phi) : \phi(0) = y, \phi(T) = z, T < \infty \right\}.$$

We say that U is a *subsolution* if for all y, z, $U(y) - U(z) \leq \mathcal{J}(y,z)$.

Statement of Main Results

Let

$$\mathcal{J}(y,z) = \inf \{ I_T(\phi) : \phi(0) = y, \phi(T) = z, T < \infty \}.$$

We say that U is a *subsolution* if for all y, z, $U(y) - U(z) \leq \mathcal{J}(y,z)$. We assume some regularity of F, and that for any compact κ there is $\alpha > 0$ such that

$$\sup_{x \in \kappa, n \in \mathbb{N}} E_x e^{\alpha \tau^n / n} < \infty.$$

Statement of Main Results

Let

$$\mathcal{J}(y,z) = \inf \left\{ I_T(\phi) : \phi(0) = y, \phi(T) = z, T < \infty \right\}.$$

We say that U is a *subsolution* if for all y, z, $U(y) - U(z) \leq \mathcal{J}(y,z)$. We assume some regularity of F, and that for any compact κ there is $\alpha > 0$ such that

$$\sup_{x \in \kappa, n \in \mathbb{N}} E_x e^{\alpha \tau^n / n} < \infty.$$

Then

- U being a subsolution is a *necessary* and *sufficient* condition for subexponential growth in number of particles and total computational effort.

Statement of Main Results

Let

$$\mathcal{J}(y,z) = \inf\left\{ I_T(\phi) : \phi(0) = y, \phi(T) = z, T < \infty \right\}.$$

We say that U is a *subsolution* if for all y, z, $U(y) - U(z) \leq \mathcal{J}(y,z)$. We assume some regularity of F, and that for any compact κ there is $\alpha > 0$ such that

$$\sup_{x\in\kappa, n\in\mathbb{N}} E_x e^{\alpha\tau^n/n} < \infty.$$

Then

- U being a subsolution is a *necessary* and *sufficient* condition for subexponential growth in number of particles and total computational effort.
- If U is a subsolution

$$\lim_{n\to\infty} -\frac{1}{n}\log E\left[(\theta^n)^2\right] = \inf_y \left\{ \mathcal{J}(x,y) + (U(x) - U(y)) \vee 0 + 2F(y) \right\}$$

Statement of Main Results

Asymptotic rate of decay:

$$\inf_y \{\mathcal{J}(x,y) + (U(x) - U(y)) \vee 0 + 2F(y)\}.$$

Statement of Main Results

Asymptotic rate of decay:

$$\inf_y \{\mathcal{J}(x,y) + (U(x) - U(y)) \vee 0 + 2F(y)\}.$$

Compare to best possible rate of decay:

$$\gamma(x) = 2\inf_y \{\mathcal{J}(x,y) + 2F(y)\}.$$

Statement of Main Results

Asymptotic rate of decay:

$$\inf_y \left\{ \mathcal{J}(x,y) + (U(x) - U(y)) \vee 0 + 2F(y) \right\}.$$

Compare to best possible rate of decay:

$$\gamma(x) = 2 \inf_y \left\{ \mathcal{J}(x,y) + 2F(y) \right\}.$$

Achieved, and hence asymptotic optimality, if at minimizing y

$$U(x) - U(y) = \mathcal{J}(x,y).$$

The F-V Quasipotential and Subsolutions

An important example.

An important example. In the context of hitting probabilities, let $A = \{x^*\}$ be stable point.

An important example. In the context of hitting probabilities, let $A = \{x^*\}$ be stable point. Define the *quasipotential*

$$Q(y) = \inf \{I_T(\phi) : \phi(T) = y, T < \infty, \phi(0) = x^*\}.$$

The F-V Quasipotential and Subsolutions

An important example. In the context of hitting probabilities, let $A = \{x^*\}$ be stable point. Define the *quasipotential*

$$Q(y) = \inf\left\{I_T(\phi) : \phi(T) = y, T < \infty, \phi(0) = x^*\right\}.$$

Then $U(y) = -Q(y)$ is always subsolution with optimal value.

The F-V Quasipotential and Subsolutions

An important example. In the context of hitting probabilities, let $A = \{x^*\}$ be stable point. Define the *quasipotential*

$$Q(y) = \inf \left\{ I_T(\phi) : \phi(T) = y, T < \infty, \phi(0) = x^* \right\}.$$

Then $U(y) = -Q(y)$ is always subsolution with optimal value.

A special case.

The F-V Quasipotential and Subsolutions

An important example. In the context of hitting probabilities, let $A = \{x^*\}$ be stable point. Define the *quasipotential*

$$Q(y) = \inf\left\{I_T(\phi) : \phi(T) = y, T < \infty, \phi(0) = x^*\right\}.$$

Then $U(y) = -Q(y)$ is always subsolution with optimal value.

A special case. Product form or asymptotically product form stochastic networks, $Q(y) = \langle a, y \rangle$.

Example, Tandem Queue

$$U(y) = -\left[\log\left(\frac{\mu_1}{\lambda}\right)\right] y_1 - \left[\log\left(\frac{\mu_2}{\lambda}\right)\right] y_2,$$

$\lambda = 1, \mu_1 = \mu_2 = 4.5.$

Example, Tandem Queue

$$U(y) = -\left[\log\left(\frac{\mu_1}{\lambda}\right)\right] y_1 - \left[\log\left(\frac{\mu_2}{\lambda}\right)\right] y_2,$$

$\lambda = 1, \mu_1 = \mu_2 = 4.5.$
Shared buffer capacity n:

n	30	40	50
Theoretical Value	2.63×10^{-18}	1.03×10^{-24}	3.80×10^{-31}
Estimate	2.63×10^{-18}	1.06×10^{-24}	3.83×10^{-31}
Std. Err.	0.08×10^{-18}	0.04×10^{-24}	0.15×10^{-31}
95% C.I.	$[2.47, 2.79] \times 10^{-18}$	$[0.99, 1.14] \times 10^{-24}$	$[3.54, 4.13] \times 10^{-31}$
Time Taken (s)	3	6	8

Example, Tandem Queue

$$U(y) = -\left[\log\left(\frac{\mu_1}{\lambda}\right)\right] y_1 - \left[\log\left(\frac{\mu_2}{\lambda}\right)\right] y_2,$$

$\lambda = 1, \mu_1 = \mu_2 = 4.5.$

Shared buffer capacity n:

n	30	40	50
Theoretical Value	2.63×10^{-18}	1.03×10^{-24}	3.80×10^{-31}
Estimate	2.63×10^{-18}	1.06×10^{-24}	3.83×10^{-31}
Std. Err.	0.08×10^{-18}	0.04×10^{-24}	0.15×10^{-31}
95% C.I.	$[2.47, 2.79] \times 10^{-18}$	$[0.99, 1.14] \times 10^{-24}$	$[3.54, 4.13] \times 10^{-31}$
Time Taken (s)	3	6	8

Separate buffers each of capacity n:

n	10	20	30
Theoretical Value	9.64×10^{-8}	1.60×10^{-15}	2.64×10^{-23}
Estimate	9.70×10^{-8}	1.57×10^{-15}	2.64×10^{-23}
Std. Err.	0.16×10^{-8}	0.03×10^{-15}	0.06×10^{-23}
95% C.I.	$[9.39, 10.0] \times 10^{-8}$	$[1.51, 1.63] \times 10^{-15}$	$[2.53, 2.75] \times 10^{-23}$
Time Taken (s)	3	12	26

Remarks

- There are ways to link the subsolution to n, improve efficiency while maintaining asymptotic optimality.

Remarks

- There are ways to link the subsolution to n, improve efficiency while maintaining asymptotic optimality.
- There is an analogous theory for importance sampling, but it imposes stronger conditions on the subsolution. Differences may be significant.