Rare-event Simulation of Brownian Motion Avoiding Hard Obstacles

Jose Blanchet (joint work with Paul Dupuis)

Columbia IEOR Department

Rubinstein’s Celebration
Agenda

- **Introduction: Rare-event Simulation**
- Brownian Motion Avoiding Obstacles
- Explaining the Strategy
- Conclusions
Assumptions:

- Let \( B(t) : t \geq 0 \) be a Brownian motion in \( \mathbb{R}^d \).
- Let \( N(t) : t \geq 0 \) be a Poisson spatial process in \( \mathbb{R}^d \) independent of \( B(t) \).

To each point of \( N(t) \) attach a square of volume \( a \) corresponding to an obstacle (obstacles can have intersections, that's fine).

Let \( T = 1 \text{st time } B(t) \text{ hits the boundary of an obstacle} \).

**Question:** Design an efficient simulation algorithm to estimate \( P(T > t) \) for large \( t \).
Assumptions:

Let \( B(t) : t \geq 0 \) be a Brownian motion in \( \mathbb{R}^d \)

Let \( (N(t) : t \geq 0) \) be a Poisson spatial process in \( \mathbb{R}^d \) independent of \( B(t) \). To each point of \( N(t) \) attach a square of volume \( a \) corresponding to an obstacle (obstacles can have intersections, that's not the case).

Let \( T = 1 \)st time \( B(t) \) hits the boundary of an obstacle

Question:

Design an efficient simulation algorithm to estimate \( \mathbb{P}(T > t) \) for large \( t \)
Assumptions:

- Let $(B(t) : t \geq 0)$ be a Brownian motion in $\mathbb{R}^d$.
- Let $(N(t) : t \geq 0)$ be a Poisson spatial process in $\mathbb{R}^d$ independent of $B(\cdot)$. 

Question:

Design an efficient simulation algorithm to estimate $\mathbb{P}(T > t)$ for large $t$. 
Assumptions:

- Let \((B(t) : t \geq 0)\) be a Brownian motion in \(R^d\).
- Let \((N(t) : t \geq 0)\) be a Poisson spatial process in \(R^d\) independent of \(B(\cdot)\).
- To each point of \(N(\cdot)\) attach a **square of volume** \(a\) corresponding to an obstacle (obstacles can have intersections, that’s fine).
Assumptions:

- Let \((B(t) : t \geq 0)\) be a Brownian motion in \(R^d\).
- Let \((N(t) : t \geq 0)\) be a Poisson spatial process in \(R^d\) independent of \(B(\cdot)\).
- To each point of \(N(\cdot)\) attach a **square of volume** \(a\) corresponding to an obstacle (obstacles can have intersections, that’s fine).
- Let \(T = 1\text{st time } B(\cdot)\) hits the boundary of an obstacle.
Assumptions:

Let $(B(t) : t \geq 0)$ be a Brownian motion in $R^d$

Let $(N(t) : t \geq 0)$ be a Poisson spatial process in $R^d$ independent of $B(\cdot)$.

To each point of $N(\cdot)$ attach a square of volume $a$ corresponding to an obstacle (obstacles can have intersections, that’s fine).

Let $T = 1$st time $B(\cdot)$ hits the boundary of an obstacle

Question:
Assumptions:

- Let \((B(t) : t \geq 0)\) be a Brownian motion in \(\mathbb{R}^d\).
- Let \((N(t) : t \geq 0)\) be a Poisson spatial process in \(\mathbb{R}^d\) independent of \(B(\cdot)\).
- To each point of \(N(\cdot)\) attach a square of volume \(a\) corresponding to an obstacle (obstacles can have intersections, that’s fine).
- Let \(T = 1st\ time\ B(\cdot)\) hits the boundary of an obstacle.

Question:

Design an efficient simulation algorithm to estimate \(P(T > t)\) for large \(t\).
Motivation

- Undetected objects for long period of time
- Motivation as a problem in random media (study of polymers in random environments)
- Materials properties (obstacles represent impurities)
- Introduced by Smoluchowsky (1918) in Chemistry and Physics / now proposed as model of molecules in motion in cells (http://jb.asm.org/cgi/content/full/187/1/23 )
- *It provides an interesting example of importance sampling that involves infinite dimensional simulation (control) problem...*. 

Blanchet (Columbia)  
Brownian Motion Avoiding Hard Obstacles  
07/08 4 / 34
Suppose want to estimate \( P(Z \in A) \)
Basic Facts about Importance Sampling

- Suppose want to estimate $P(Z \in A)$
- Importance sampling estimation says: Find an appropriate change-of-measure $Q(d\omega)$ and produce the importance sampling (IS) estimator

$$Y = \frac{dP}{dQ}(\omega) I(Z(\omega) \in A)$$
Suppose want to estimate $P(Z \in A)$

Importance sampling estimation says: Find an appropriate change-of-measure $Q(d\omega)$ and produce the importance sampling (IS) estimator

$$Y = \frac{dP}{dQ}(\omega) I(Z(\omega) \in A)$$

Simulate iid replications of $Y$ to estimate $P(Z \in A) = E^Q Y$. 

Basic Facts about Importance Sampling

- **Suppose want to estimate** $P(Z \in A)$
- Importance sampling estimation says: Find an appropriate *change-of-measure* $Q(d\omega)$ and produce the importance sampling (IS) estimator
  \[ Y = \frac{dP}{dQ}(\omega) I(Z(\omega) \in A) \]
- Simulate iid replications of $Y$ to estimate $P(Z \in A) = E^Q Y$.
- Want to reduce the variance of $Y$
An Obvious Observation and a Powerful Principle

- Select $Q(\cdot)$ as conditional distribution given $Z \in A$

$$Q(d\omega) = \frac{I(Z(\omega) \in A) P(d\omega)}{P(Z \in A)},$$

Obviously useless to implement but yields a powerful principle: WE SHALL CALL IT GISP ("Good Importance Sampling Principle").

GISP: "To design a good importance sampling try to mimic the conditional distribution of the process given the rare event" (Asmussen and Rubinstein '85)
Select $Q(\cdot)$ as conditional distribution given $Z \in A$

$$Q(d\omega) = \frac{I(Z(\omega) \in A) P(d\omega)}{P(Z \in A)},$$

Then $Y = P(Z \in A)$ is unbiased with zero variance...
An Obvious Observation and a Powerful Principle

- Select $Q(\cdot)$ as conditional distribution given $Z \in A$

$$Q(d\omega) = \frac{I(Z(\omega) \in A) P(d\omega)}{P(Z \in A)},$$

- Then $Y = P(Z \in A)$ is unbiased with zero variance...

- Obviously useless to implement BUT yields a powerful principle: WE SHALL CALL IT GISP ("Good Importance Sampling Principle").
An Obvious Observation and a Powerful Principle

- Select $Q(\cdot)$ as conditional distribution given $Z \in A$

\[ Q(d\omega) = \frac{I(Z(\omega) \in A) P(d\omega)}{P(Z \in A)} , \]

- Then $Y = P(Z \in A)$ is unbiased with zero variance...

- Obviously useless to implement BUT yields a powerful principle: WE SHALL CALL IT GISP ("Good Importance Sampling Principle").

- **GISP**: "To design a good importance sampling try to mimic the conditional distribution of the process given the rare event" (Asmussen and Rubinstein ’85)
Goal of GISP: Finding an efficient or *asymptotically optimal* estimator.
Goal of GISP: Finding an efficient or asymptotically optimal estimator

Definition: Given $\alpha_n = P(A_n) \to 0$ as $n \to \infty$ we say that $Z_n$ is asymptotically optimal or (weakly) efficient if $\alpha_n = EZ_n$ and

$$\lim_{n \to \infty} \frac{\log EZ_n^2}{\log \alpha_n} = 2.$$
Goal of GISP: Finding an efficient or *asymptotically optimal estimator*

Definition: Given $\alpha_n = P(A_n) \to 0$ as $n \to \infty$ we say that $Z_n$ is *asymptotically optimal* or (weakly) efficient if $\alpha_n = EZ_n$ and

$$\lim_{n \to \infty} \frac{\log EZ_n^2}{\log \alpha_n} = 2.$$ 

Remark: Need to also consider the computer time to generate $Z_n$ that is typically polynomial in $|\log \alpha_n|$ so doesn’t contribute significantly to complexity.
Agenda

- Introduction: Rare-event Simulation
- **Brownian Motion Avoiding Obstacles**
- Explaining the Strategy
- Conclusions
Recall $T = 1$st time $B(\cdot)$ hits a Poissonian obstacle

$$\begin{align*}
P ( T > t ) &= E \left( P ( T > t | B ( s ) : 0 \leq s \leq t ) \right) \\
&= E \left( P \left( \text{No obstacle in trajectory} | B ( \cdot ) \right) \right) \\
&= E \exp (-V(t, a)),
\end{align*}$$

where

$$V(t, a) = Vol \left( \bigcup_{0 \leq s \leq t} \text{Square (center} = B(s), \text{vol} = a) \right)$$
By the invariance principle, if \( \tau = t^{d/(d+2)} \) then we have

\[
E \exp (-V(t,a)) = E \exp \left(-\tau V \left( \tau, a\tau^{-1/d} \right) \right)
\]
By the invariance principle, if $\tau = t^{d/(d+2)}$ then we have

$$E \exp(-V(t,a)) = E \exp(-\tau V(\tau, a\tau^{-1/d}))$$

We define and study estimation for

$$\alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta))$$
By the invariance principle, if $\tau = t^{d/(d+2)}$ then we have

$$E \exp (-V(t,a)) = E \exp \left(-\tau V \left(\tau, a\tau^{-1/d}\right) \right)$$

We define and study estimation for

$$\alpha(\tau, \delta) = E_0 \exp (-\tau V(\tau, \delta))$$

How to obtain a "GISP" here? What does large deviations tell us?
Donsker and Varadhan '75 proved that when $\delta \downarrow 0 \geq a\tau^{-1/d}$ as $\tau \downarrow 0$ (also Bolthausen '90, Sznitman '89) then

$$\frac{1}{\tau} \log \alpha(\tau, \delta) \to -\inf_{f: \int f = 1} \left( \text{vol}(\text{supp}(f)) + \frac{1}{8} \int \frac{\|\nabla f\|^2}{f} \right)$$

$$= -\inf_{G \text{ open}} (\text{vol}(G) + \lambda_G),$$

where $\lambda_G = \text{principal e-value of } \triangle/2$ on $G \to$ discuss optimal path
Conditional description (Schmock $d = 1$, Sznitman $d = 2$, Povel $d > 2$): B. Motion travels $O\left(\tau^{1/d}\right)$ distance to find an optimal center (random even at $\tau^{1/d}$ scales!) and it confines itself inside a ball with optimal radius at spatial scales of $O\left(\tau^{1/d}\right)$. 

Brownian motion in 2 dimensions

Picture at scale of order $O(\tau^{1/d})$
Question:
Summary: What is the Problem?

- Question:
  * How to describe a change-of-measure that mimics the conditional distribution close enough to obtain an asymptotically optimal estimator – GISP?
Question:

How to describe a change-of-measure that mimics the conditional distribution close enough to obtain an asymptotically optimal estimator – GISP?

Such change-of-measure must find an optimal ball with the right distribution and do it step-by-step from the Brownian path...
Agenda

- Introduction: Rare-event Simulation
- Brownian Motion Avoiding Obstacles
- **Explaining the Strategy**
- Conclusions
The Strategy

- $\alpha (\tau, \delta) = E_0 \exp (-\tau V (\tau, \delta))$
- Divide the space in cubes of volume

![Diagram showing a grid with a yellow spot indicating an obstacle of area $\delta$.]
We generate a *suitable process* that keeps exploring regions as follows:

Initial “explored” region.
Explored regions are painted pink.
The Strategy

- The distribution of the process adapts according to explored regions (we’ll see how!)
The Strategy

Blanchet (Columbia)
Brownian Motion Avoiding Hard Obstacles

07/08 18 / 34
And one goes on sequentially — now we’ll explain the evolution
Recall the goal: $\alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta))$
The Strategy

- Recall the goal: $\alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta))$
- Given total pink region $\mathcal{R}_M$ (say $R_0 \cup R_1 \cup ... \cup R_M$), where $M$ is the region \textit{JUST visited}
Recall the goal: \( \alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta)) \)

Given total pink region \( R_M \) (say \( R_0 \cup R_1 \cup \ldots \cup R_M \)), where \( M \) is the region \textit{JUST visited}

Want to spend as much as possible in the explored region (which is free of obstacles!)
The Strategy

- Recall the goal: \( \alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta)) \)
- Given total pink region \( \mathcal{R}_M \) (say \( R_0 \cup R_1 \cup \ldots \cup R_M \)), where \( M \) is the region *JUST visited*
- Want to spend as much as possible in the explored region (which is free of obstacles!)
- Let \( T_M = \inf\{t \geq 0 : B(t) \notin \mathcal{R}_M\} \ldots \)
The Strategy

- Recall the goal: $\alpha(\tau, \delta) = E_0 \exp(-\tau V(\tau, \delta))$
- Given total pink region $\mathcal{R}_M$ (say $R_0 \cup R_1 \cup \ldots \cup R_M$), where $M$ is the region *JUST visited*
- Want to spend as much as possible in the explored region (which is free of obstacles!)
- Let $T_M = \inf\{t \geq 0 : B(t) \notin \mathcal{R}_M\}$...
- Select $\theta_M$ such that
  
  $$E(\exp(\theta_M T_M) | \text{Visited region } \mathcal{R}_M) = \exp(\gamma \varepsilon \tau),$$

  **AND** $\gamma$ which will be chosen...
The Strategy

- Implement the strategy sequentially: Given region $\mathcal{R}_M$ sample according to the SDE

$$dX(t) = \nabla \log \nu_{\mathcal{R}_M}(X(t), \theta_M) \, dt + dB(t),$$

where $\nu_{\mathcal{R}_M}(x) = E_x(\exp(\theta_M T_M)|\text{Visited region } \mathcal{R}_M)$ for $x \in \mathcal{R}_M$. 
The Strategy

- Implement the strategy sequentially: Given region $\mathcal{R}_M$ sample according to the SDE

$$dX(t) = \nabla \log v_{\mathcal{R}_M}(X(t), \theta_M) \, dt + dB(t),$$

where $v_{\mathcal{R}_M}(x) = E_x(\exp(\theta_M T_M) | \text{Visited region } \mathcal{R}_M)$ for $x \in \mathcal{R}_M$.

- The likelihood ratio

$$L_\tau = \frac{1}{v_{\mathcal{R}_M(\tau)}(B_\tau, \theta_M(\tau))} \exp \left( \gamma \varepsilon \tau M_\tau - \int_0^\tau \theta_M(s) \, ds \right),$$
So, the I.S. estimator is

\[
L_\tau \exp(-\tau V(\tau, \delta)) = \frac{\exp\left(-\tau (V(\tau, \delta) - \gamma \epsilon M_\tau) - \int_0^\tau \theta_M(s) \, ds\right)}{\nu_{R_M(\tau)}(B_\tau, \theta_M(\tau))},
\]

So, the I.S. estimator is

\[ L_\tau \exp\left(-\tau V(\tau, \delta)\right) = \frac{\exp\left(-\tau \left(V(\tau, \delta) - \gamma \varepsilon M_\tau\right) - \int_0^\tau \theta_{M(s)} ds\right)}{\nu_{R_{M(\tau)}}(B_\tau, \theta_{M(\tau)})}, \]

\[ \nu_{R_{M}}(x, \theta_{M}) = E_x\left(\exp\left(\theta_{M} T_{M}\right) | R_{M}\right) \geq 1 \]
So, the I.S. estimator is

\[ L_\tau \exp(-\tau V(\tau, \delta)) \leq \exp\left(-\tau (V(\tau, \delta) - \gamma \varepsilon M_{\tau}) - \int_0^\tau \theta_{M(s)} ds \right), \]

\[ v_{R_M}(x, \theta_M) = E_x(\exp(\theta_M T_M) | R_M) \geq 1 \]
So, the I.S. estimator is

$$L_{\tau} \exp(-\tau V(\tau, \delta)) \leq \exp\left(-\tau (V(\tau, \delta) - \gamma \varepsilon M_\tau) - \int_0^\tau \theta_M(s) \, ds\right),$$

1. $v_{\mathcal{R}_M}(x, \theta_M) = E_x(\exp(\theta_M T_M) | \mathcal{R}_M) \geq 1$
2. $P(T_M > x | \mathcal{R}_M) = \exp(-\lambda_{\mathcal{R}_M} x + o(x))$
3. **By the choice of $\theta_M$, we have that** $\theta_M(s) = \lambda_{\mathcal{R}_M(s)} + o(1/\tau)$
So, the I.S. estimator is

\[ L_\tau \exp(-\tau V(\tau, \delta)) \leq \exp\left(-\tau (V(\tau, \delta) - \gamma \varepsilon M_\tau) - \int_0^\tau \lambda_{\mathcal{R}_M(s)} ds\right), \]

1. \( \nu_{\mathcal{R}_M}(x, \theta_M) = E_x(\exp(\theta_M T_M)|\mathcal{R}_M) \geq 1 \)
2. \( P(T_M > x|\mathcal{R}_M) = \exp(-\lambda_{\mathcal{R}_M} x + o(x)) \)
3. By the choice of \( \theta_M \), we have that \( \theta_M(s) = \lambda_{\mathcal{R}_M(s)} + o(1/\tau) \)
The Strategy

So, the I.S. estimator is

$$L_\tau \exp \left( -\tau V (\tau, \delta) \right) \leq \exp \left( -\tau \left( V (\tau, \delta) - \gamma \varepsilon M_\tau \right) - \int_0^\tau \lambda R_M(s) \, ds \right),$$

1. \( \nu_{R_M} (x, \theta_M) = E_x (\exp (\theta_M T_M) | R_M) \geq 1 \)
2. \( P ( T_M > x | R_M) = \exp (-\lambda R_M x + o(x)) \)
3. By the choice of \( \theta_M \), we have that \( \theta_M(s) = \lambda R_M(s) + o(1/\tau) \)
4. \( \lambda R_M(s) \geq \lambda R_M(\tau) \geq 0 \text{ for } r \geq s \)
So, the I.S. estimator is

\[ L_\tau \exp (-\tau V(\tau, \delta)) \leq \exp \left( -\tau \left( V(\tau, \delta) - \gamma \varepsilon M_\tau \right) - \tau \lambda_{R_M(\tau)} \right), \]

1. \( v_{R_M}(x, \theta_M) = E_x(\exp(\theta_M T_M) | R_M) \geq 1 \)
2. \( P(T_M > x | R_M) = \exp(-\lambda_{R_M} x + o(x)) \)
3. By the choice of \( \theta_M \), we have that \( \theta_M(s) = \lambda_{R_M(s)} + o(1/\tau) \)
4. \( \lambda_{R_M(s)} \geq \lambda_{R_M(\tau)} \geq 0 \) for \( r \geq s \)
So, the I.S. estimator is

\[ L_\tau \exp \left( -\tau V(\tau, \delta) \right) \leq \exp \left( -\tau (V(\tau, \delta) - \gamma \varepsilon M_\tau) - \tau \lambda_{\mathcal{R}_{M(\tau)}} \right), \]

1. \( v_{\mathcal{R}_M}(x, \theta_M) = E_x(\exp(\theta_M T_M) | \mathcal{R}_M) \geq 1 \)
2. \( P(T_M > x | \mathcal{R}_M) = \exp(-\lambda_{\mathcal{R}_M} x + o(x)) \)
3. By the choice of \( \theta_M \), we have that \( \theta_M(s) = \lambda_{\mathcal{R}_M(s)} + o(1/\tau) \)
4. \( \lambda_{\mathcal{R}_M(s)} \geq \lambda_{\mathcal{R}_M(\tau)} \geq 0 \) for \( r \geq s \)
5. \( \varepsilon M_\tau = \text{Vol}(\mathcal{R}_{M(\tau)}) \leq V(\tau, \delta) \) (for \( \varepsilon \leq \delta/2 \))
The Strategy

So, the I.S. estimator is

\[ L_\tau \exp (-\tau V(\tau, \delta)) \leq \exp \left( -\tau \left( (1 - \gamma) \text{Vol}(R_{M(\tau)}) + \lambda R_{M(\tau)} \right) \right) \]

1. \( v_{R_M}(x, \theta_M) = E_x \left( \exp (\theta_M T_M) | R_M \right) \geq 1 \)
2. \( P(T_M > x | R_M) = \exp (-\lambda_{R_M} x + o(x)) \)
3. By the choice of \( \theta_M \), we have that \( \theta_M(s) = \lambda_{R_M(s)} + o(1/\tau) \)
4. \( \lambda_{R_M(s)} \geq \lambda_{R_M(\tau)} \geq 0 \) for \( r \geq s \)
5. \( \epsilon M_\tau = \text{Vol}(R_{M(\tau)}) \leq V(\tau, \delta) \). For \( \epsilon \leq \delta/2 \)
The Strategy

- Take $\varepsilon = \delta$ and this takes us to...

$$\text{IS EST}$$

$$\leq \exp \left( -\tau \min_{G: \text{open}} (\text{Vol}(G) (1 - \gamma) + \lambda_G) \right)$$

$$= \exp (\tau O(\gamma)) \exp \left( -\tau \min_{G: \text{open}} (\text{Vol}(G) + \lambda_G) \right)$$
The Strategy

- Take $\varepsilon = \delta$ and this takes us to...

\[
\text{IS EST} \\
\leq \exp \left( -\tau \min_{G: \text{open}} (\text{Vol}(G)(1 - \gamma) + \lambda_G) \right) \\
= \exp(\tau O(\gamma)) \exp \left( -\tau \min_{G: \text{open}} (\text{Vol}(G) + \lambda_G) \right)
\]

- Therefore

\[
E(\text{Estimator}^2) = \alpha(\tau, \delta)^2 \exp(\tau O(\gamma) + o(\tau))
\]

and weak efficiency follows (for $\delta, \gamma \to 0$ sufficiently slow as $\tau \to \infty$)
A path for $\tau = 1000$, using $\gamma = .1$
A path for $\tau = 5000$, using $\gamma = .1$
Agenda

- Introduction: Rare-event Simulation
- Brownian Motion Avoiding Obstacles
- Explaining the Strategy
- Conclusions
Brownian motion avoiding obstacles gives an example where history dependent importance sampling should be performed to achieve efficiency.

Strategy induces confinement $\rightarrow$ particle tries to stay inside explored region, which is obstacle free.

Eventually, explored region is basically a ball with optimal radius and specific distribution center.