Exact simulation of the first-passage time of diffusions

S. Herrmann

University of Burgundy, Dijon

joint work with Cristina ZUCCA, University of Torino, Italy

June 25ᶜ 2018

Outline

Introduction to the first-passage time

Modeling biological or physical stochastic systems often requires to handle with one-dimensional diffusion processes.

Two types of information:

1. the marginal probability distribution function at a fixed time t.
2. the description of the whole paths.

Introduction to the first-passage time

Modeling biological or physical stochastic systems often requires to handle with one-dimensional diffusion processes.

Two types of information:

1. the marginal probability distribution function at a fixed time t.
2. the description of the whole paths.

The marginal pdf is insufficient in many applications:

- financial derivatives with barriers
- ruin probability of an insurance fund
- optimal stopping problems
- neuronal sciences

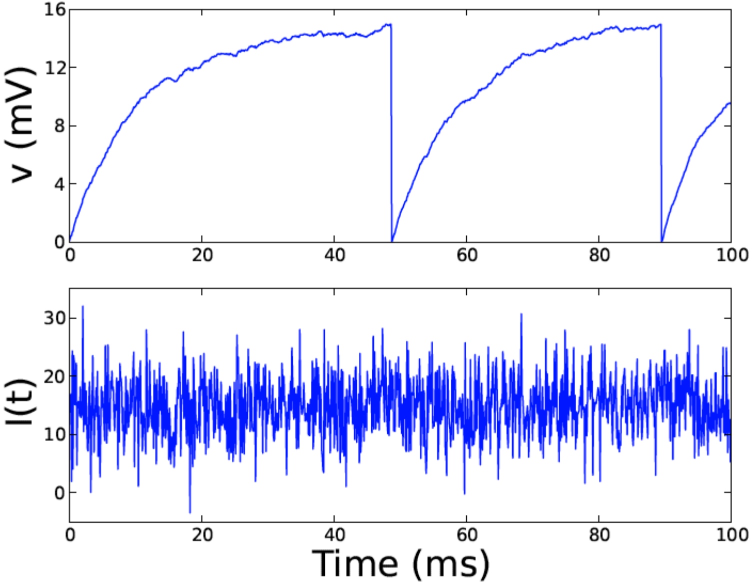
Some Integrate and Fire models define the *spiking times* as the first hitting time of a threshold by the membrane potential. If the membrane potential is given by a stochastic differential equation, the spiking times are the **first hitting times of the threshold by such a diffusion.**

The leaky integrate-and-fire (LIF) neuron is probably one of the simplest spiking neuron models, its input signal is given by $I(t)$:

$$\pi_m \frac{dv(t)}{dt} = -v(t) + R\ I(t)$$

- $v(t)$ represents the membrane potential at time t,
- π_m is the membrane time constant
- R is the membrane resistance.

When the membrane potential $v(t)$ reaches a threshold v^{th} (spiking threshold), it is instantaneously reset to a lower value v^r (reset potential) and the leaky integration process starts anew with the initial value v^r.


v (mV)
16
12
8
4
0
0
20
40
60
80
100
I(t)
30
20
10
0
0
20
40
60
80
100
Time (ms)

First-passage time τ_L

Let $(X_t,\ t \geq 0)$ be a one-dimensional diffusion process satisfying

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \quad X_0 = x < L.$$

Aim: simulation of the FPT defined by $\tau_L := \inf\{t \geq 0 : X_t = L\}$.

Different tools for simulation purposes: explicit expression of the pdf, approximation of the stochastic process, rejection sampling...

First-passage time τ_L

Let $(X_t,\ t \geq 0)$ be a one-dimensional diffusion process satisfying

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \quad X_0 = x < L.$$

Aim: simulation of the FPT defined by $\tau_L := \inf\{t \geq 0 : X_t = L\}$.

Different tools for simulation purposes: explicit expression of the pdf, approximation of the stochastic process, rejection sampling...

Standard Brownian case ($B_0 = 0$):

The optional stopping thm applied to $M_t = \exp\{\lambda B_t - \frac{1}{2}\lambda^2 t\}$ leads to

$$\mathbb{E}[e^{-\lambda\tau_L}] = e^{-\sqrt{2\lambda}L}, \quad \lambda \geq 0.$$

Inversion of the Laplace transform:

$$\mathbb{P}(\tau_L \in dt) = \frac{1}{\sqrt{2\pi t^3}}\ e^{-\frac{L^2}{2t}}dt, \quad t > 0.$$

Hence $\tau_L \sim L^2/G^2$

where $G \sim \mathcal{N}(0,1)$.

Easy and exact simulation !

General one-dimensional diffusion processes:

We define the generator associated to the diffusion $(X_t,\ t \geq 0)$ by

$$Lf(x) = \frac{\sigma^2(x)}{2}\frac{d^2 f}{dx^2}(x) + b(x)\frac{df}{dx}(x), \quad \text{for } x \in \mathbb{R}.$$

Then the Laplace transform of the FPT is the unique solution of the following Sturm-Liouville boundary value problem on $]-\infty, L[$:

$$\begin{cases} Lu(x) = \lambda u(x), \\ u\,\big|_{x=L} = 1 \\ \lim_{x\to-\infty} u(x) = 0. \end{cases}$$

Let ψ_λ the unique increasing positive solution of $Lu = \lambda u$.

The following property holds:

$$\mathbb{E}_x[e^{-\lambda\tau_L}] = \frac{\psi_\lambda(x)}{\psi_\lambda(L)}$$

General one-dimensional diffusion processes:

We define the generator associated to the diffusion $(X_t,\ t \geq 0)$ by

$$Lf(x) = \frac{\sigma^2(x)}{2}\frac{d^2 f}{dx^2}(x) + b(x)\frac{df}{dx}(x), \quad \text{for } x \in \mathbb{R}.$$

Then the Laplace transform of the FPT is the unique solution of the following Sturm-Liouville boundary value problem on $]-\infty, L[$:

$$\begin{cases} Lu(x) = \lambda u(x), \\ u\,|_{x=L} = 1 \\ \lim_{x\to-\infty} u(x) = 0. \end{cases}$$

Let ψ_λ the unique increasing positive solution of $Lu = \lambda u$.

The following property holds:

$$\mathbb{E}_x[e^{-\lambda\tau_L}] = \frac{\psi_\lambda(x)}{\psi_\lambda(L)} = \frac{\mathcal{H}_{-\lambda/\theta}(x\sqrt{\theta})}{\mathcal{H}_{-\lambda/\theta}(L\sqrt{\theta})}.$$

- Ornstein-Uhlenbeck case $(\sigma = 1,\ b(x) = -\theta x)$: *Hermite* functions

$$\mathcal{H}_\nu(z) = \frac{1}{2\Gamma(-\nu)}\sum_{m\geq 0}\frac{(-1)^m}{m!}\Gamma\Big(\frac{m-\nu}{2}\Big)(2z)^m.$$

- When the transition probability of (X_t) has an explicit expression...

Let us define

$$\begin{cases} f(t,x|s,y)dx := \mathbb{P}(X_t \in dx | X_s = y), \quad s \leq t, \\ \varphi(t,x|s,y) = b(x)f(t,x|s,y) - \frac{1}{2}\frac{\partial}{\partial x}\Big[\sigma^2(x)f(t,x|s,y)\Big]. \end{cases}$$

φ represents the probability current of the diffusion process.

Voltera-type integral equation (see Buonocore, Nobile, Ricciardi)

The pdf $f_L(t)$ of the FPT τ_L satisfies the Voltera-type equation:

$$f_L(t) = 2\varphi(L,t|x,0) - 2\int_0^t f_L(s)\varphi(L,t|L,s)ds, \quad \text{with } X_0 = x.$$

- When the transition probability of (X_t) has an explicit expression...

Let us define

$$\begin{cases} f(t,x|s,y)dx := \mathbb{P}(X_t \in dx | X_s = y), \quad s \leq t, \\ \varphi(t,x|s,y) = b(x)f(t,x|s,y) - \frac{1}{2}\frac{\partial}{\partial x}\Big[\sigma^2(x)f(t,x|s,y)\Big]. \end{cases}$$

φ represents the probability current of the diffusion process.

Voltera-type integral equation (see Buonocore, Nobile, Ricciardi)

The pdf $f_L(t)$ of the FPT τ_L satisfies the Voltera-type equation:

$$f_L(t) = 2\varphi(L,t|x,0) - 2\int_0^t f_L(s)\varphi(L,t|L,s)ds, \quad \text{with } X_0 = x.$$

Closed form results for the Brownian motion and for the O-U process.
In general: numerical approximation of the integral... (it works fine due to the particular choice of the Voltera kernel – non singular !)
What about the simulation of τ_L ?

- General method: time discretization

Instead of considering the approximation of the pdf, it is possible to deal directly with an approximation of the diffusion process (Euler scheme).

$$X_{(n+1)\Delta} = X_{n\Delta} + \Delta b(X_{n\Delta}) + \sqrt{\Delta}\,\sigma(X_{n\Delta})G_n, \quad n \geq 0,$$

where (G_n) stands for a sequence of independent Gaussian distributed r.v.

Let τ_L^Δ be the FPT of the **discrete-time process.**

Overestimation of the FPT: $\tau_L \underset{st}{\leq} \tau_L^\Delta$

Important to improve the algorithm:

1. a shift of the boundary (Broadie-Glasserman-Kou, Gobet-Menozzi)
2. computation of the probability for a Brownian bridge to hit the boundary during a small time interval (Giraudo-Saccerdote-Zucca)

Advantage: rough description the paths. **But:** bounded time interval !

Acceptance-rejection sampling: an exact simulation of the FPT

Principal idea: Let f and g two probability distribution functions, such that $h(x) := f(x)/g(x)$ is upper-bounded by a constant $c > 0$.

Aim: simulation of X with pdf f.

1. Generate a rv Y with pdf g.
2. Generate U uniformly distributed (independent from Y).
3. If $U \leq h(Y)/c$, then set $X = Y$; otherwise go back to 1.

Acceptance-rejection sampling: an exact simulation of the FPT

Principal idea: Let f and g two probability distribution functions, such that $h(x) := f(x)/g(x)$ is upper-bounded by a constant $c > 0$.
Aim: simulation of X with pdf f.

1. Generate a rv Y with pdf g.
2. Generate U uniformly distributed (independent from Y).
3. If $U \leq h(Y)/c$, then set $X = Y$; otherwise go back to 1.

Important: h should be bounded and have an explicit expression !
Application to the first passage problem: the Girsanov transformation permits to

- link the distribution of the diffusion process $(X_t,\ t \geq 0)$ to the Brownian one $(B_t,\ t \geq 0)$.
- give an expression of the function h.

Girsanov's transformation was already used for simulation purposes by Beskos and Roberts (exact simulation on some fixed interval $[0, T]$).

From now on, $\sigma = 1$ (diffusion coefficient). We assume that the drift term $b \in \mathcal{C}^1(]-\infty, L])$ and introduce $\beta(x) = \int_0^x b(y)dy$ and $\gamma := \frac{b^2+b'}{2}$.

Girsanov's transformation

For any bounded measurable function $\psi : \mathbb{R} \to \mathbb{R}$, we obtain

$$\mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L<\infty\}}] = \mathbb{E}_{\mathbb{Q}}[\psi(\tau_L)\eta(\tau_L)] \exp\Big\{\beta(L) - \beta(x)\Big\},$$

where $\mathbb{P}$ (resp. $\mathbb{Q}$) corresponds to X (resp. B) and

$$\eta(t) := \mathbb{E}\Big[\exp - \int_0^t \gamma(L - R_s)ds \Big| R_t = L - x\Big].$$

Here $(R_t,\ t \geq 0)$ stands for a 3-dimensional Bessel process with $R_0 = 0$.

Proof : Girsanov + Itô's formula + conditional distribution.

$$\mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L<\infty\}}] = \mathbb{E}_{\mathbb{Q}}\Big[\psi(\tau_L)\exp\left(\int_0^{\tau_L} b(B_s)dB_s - \frac{1}{2}\int_0^{\tau_L} b^2(B_s)ds\right)\Big] \quad \square$$

$$\begin{cases} \mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L<\infty\}}] = \mathbb{E}_{\mathbb{Q}}[\psi(\tau_L)\eta(\tau_L)] \exp\Big\{\beta(L)-\beta(x)\Big\}, \\ \eta(t) := \mathbb{E}\Big[\exp - \int_0^t \gamma(L-R_s)ds \Big| R_t = L-x\Big]. \end{cases}$$

Advantages:

- Under $\mathbb{Q}$, it is easy to generate τ_L.
- An appropriate situation for a rejection method, if $\tau_L < \infty$ under $\mathbb{P}$.

Difficulties:

- the boundedness of $\eta(t)$ for $t \geq 0$. We suggest in a first phase to assume: $\gamma(x) \geq 0$ for all $x \in \mathbb{R}$.
- the non-explicit expression of $\eta(t)$: we shall assume that $\gamma(x) \leq \kappa$ for all $x \in \mathbb{R}$ and introduce a Poisson Point Process.

To sum up, the main assumption becomes:

$$0 \leq \gamma(x) \leq \kappa.$$

Algorithm (A1) or (A2).

Step 1: *Simulate a r.v.* $T = (L-x)^2/G^2$ *with* $G \sim \mathcal{N}(0,1)$.

Step 2: *Simulate a 3-dimensional Bessel process* (R_t) *on the time interval* $[0, T]$ *with endpoint* $R_T = L - x$ *and define*

$$D_{R,T} := \Big\{(t,v) \in [0,T] \times \mathbb{R}_+ : \ v \le \gamma(L - R_t)\Big\}.$$

Step 3: *Simulate a Poisson point process* N *on the state space* $[0, T] \times \mathbb{R}_+$, *independent of the Bessel process, whose intensity measure is the Lebesgue one.*

Step 4: *If* $N(D_{R,T}) = 0$ *then set* $Y = T$ *otherwise go to Step 1.*

Theorem (theoretical viewpoint)

The outcome Y and the FPT of the diffusion process τ_L are identically distributed.

Efficiency of the algorithm.

Remark: Be carefull with the simulation of the PPP: if you sample all points, their averaged number is $\mathbb{E}[\kappa T] = \infty$: efficiency to be improved !

We define:

- $\mathcal{I}$ the number of iterations (step 1)
- $\mathcal{N}_1, \ldots, \mathcal{N}_\mathcal{I}$ the numbers of random points (Poisson process) used for each iteration.
- $\mathcal{N}_\Sigma = \mathcal{I} + \mathcal{N}_1 + \ldots + \mathcal{N}_\mathcal{I}$ the total number of r.v.

Proposition

The following upper-bound holds $\mathbb{E}[\mathcal{I}] \leq \exp((L - x)\sqrt{2\kappa})$.

Efficiency of the algorithm.

Remark: Be carefull with the simulation of the PPP: if you sample all points, their averaged number is $\mathbb{E}[\kappa T] = \infty$: efficiency to be improved !

We define:

- $\mathcal{I}$ the number of iterations (step 1)
- $\mathcal{N}_1, \ldots, \mathcal{N}_{\mathcal{I}}$ the numbers of random points (Poisson process) used for each iteration.
- $\mathcal{N}_\Sigma = \mathcal{I} + \mathcal{N}_1 + \ldots + \mathcal{N}_{\mathcal{I}}$ the total number of r.v.

Proposition

The following upper-bound holds $\mathbb{E}[\mathcal{I}] \leq \exp((L-x)\sqrt{2\kappa})$.

Reduction of the number of iterations:

- For $(L-x)$, linearization by space splitting.
- For κ: if $0 < \gamma_0 \leq \gamma(x) \leq \kappa$ for all $x \in \mathbb{R}$, then replace $\gamma(\cdot) \leftarrow \gamma(\cdot) - \gamma_0$, $\kappa \leftarrow \kappa - \gamma_0$ & introduce the simulation of $IG\left(\frac{L-x}{\sqrt{2\gamma_0}}, (L-x)^2\right)$ *(Michael-Schucany-Haas).*

Proposition: number of r.v. during the first iteration.

Assumption: $\exists C_\gamma > 0$, $\exists r < 1$ such that

$$\inf_{y \le z \le L} \gamma(z) \ge C_\gamma |y|^{-r}, \quad \text{for all} \quad y \le -1.$$

Then $\exists M_{\gamma,1} > 0$ and $\exists M_{\gamma,2} > 0$ s.t. the number of random points satisfies

$$\mathbb{E}[\mathcal{N}_1] \le M_{\gamma,1} + \kappa M_{\gamma,2}(x^2 + (L-x)^{(1+r)/2}), \quad \text{for } x < L.$$

Proof: $\mathbb{E}_c[\mathcal{N}_1] = H_T + \kappa I_T$ with

$$\begin{cases} H_T := e^{-\int_0^T \gamma(L-R_w)\,dw} \le 1 \\ I_T := \int_0^T e^{-\int_0^u \gamma(L-R_w)\,dw}\,du. \end{cases}$$

Bounds for the (3d)-Bessel bridge:

$$R_{sT} \underset{st}{\le} L - x + \sqrt{T}\,\overline{R}_s, \quad s \in [0,1].$$

$(\overline{R}_s)_{s\ge 0}$ is a standard Bessel bridge.

$$\mathbb{P}(\sup_{[0,1]} \overline{R}_u > T^\alpha) \le \frac{\sqrt{e\pi}}{4\sqrt{2}}\,\frac{\pi T^\alpha}{\sinh^2(T^\alpha)}.$$

Using the agreement formula (see Chung or Pitman-Yor), we obtain

$$\mathbb{P}\Big(\sup_{u\in[0,1]} \bar{R}_u > T^\alpha\Big) = C_3\mathbb{E}\Big[\sqrt{\bar{\tau}}1_{\{\bar{\tau}<T^{-2\alpha}\}}\Big].$$

Here $C_3 = \sqrt{2}/\Gamma(3/2)$ and $\bar{\tau} = \tau + \hat{\tau}$ where τ is the first hitting time of the level 1 for a 3-dimensional Bessel process and $\hat{\tau}$ an independent copy of τ.

$$\mathbb{P}\Big(\sup_{u\in[0,1]} \bar{R}_u > T^\alpha\Big) \le C_3\, T^{-\alpha}\, \mathbb{P}\Big(\exp -\lambda\bar{\tau} > \exp -\lambda T^{-2\alpha}\Big)$$

$$\le C_3\, T^{-\alpha}\, e^{\lambda T^{-2\alpha}}\mathbb{E}[e^{-\lambda\bar{\tau}}] = C_3\, T^{-\alpha}\, e^{\lambda T^{-2\alpha}}\frac{(2\lambda)^{1/2}}{C_3^2 I_{1/2}^2(\sqrt{2\lambda})},$$

for any $\lambda > 0$. I_ν stands for the Bessel function of the first kind. In particular $I_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sinh x$. The particular choice $\lambda = T^{2\alpha}/2$ leads to

$$\mathbb{P}\Big(\sup_{u\in[0,1]} \bar{R}_u > T^\alpha\Big) \le \frac{\sqrt{e\pi}}{2\sqrt{2}}\,\frac{\pi T^\alpha}{2\sinh^2(T^\alpha)}. \qquad \square$$

Examples of generalization and numerics

Example 1. $dX_t = (2 + \sin(X_t))\, dt + dB_t,\ X_0 = 0$. We have $0 \le \gamma \le 5$.

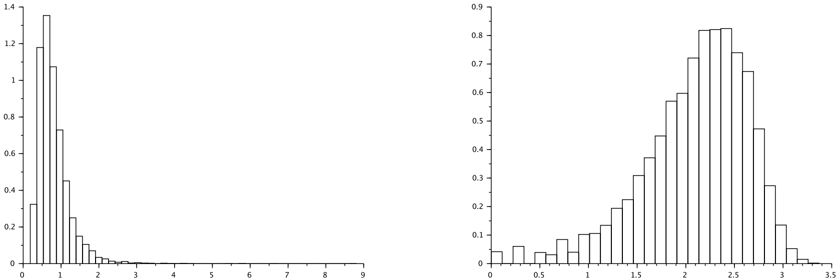
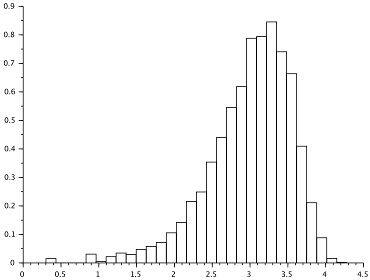


Figure: Histogram of the hitting time distribution for 10 000 simulations corresponding to the level $L = 2$ and starting position $X_0 = 0$ (left), histogram of the number of iterations in Algorithm (A1) in the $\log_{10}$-scale (right).



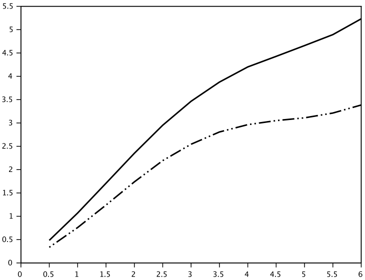


Figure: Number of random variables used in Algorithm (A1) for 10 000 simulations with $L = 2$, $X_0 = 0$ in the $\log_{10}$-scale (left) and mean number of iterations versus the level height L for Algorithm $(\mathrm{A1})_{\mathrm{shift}}$ respectively *(A1)* (dashed line resp. solid line), both curves are in the $\log_{10}$-scale (10 000 simulations have been used for the average estimation).

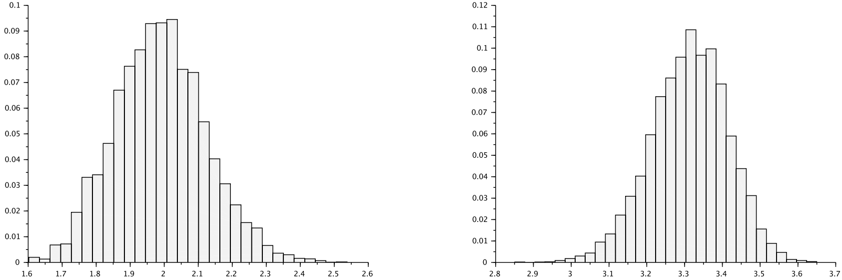


Figure: Histogram of the number of random variables in Algorithm (A1) using space splitting for 10 000 with $L = 2$, $X_0 = 0$, $k = 20$ (left), $L = 20$, $k = 20$ (right) both in the $\log_{10}$-scale.

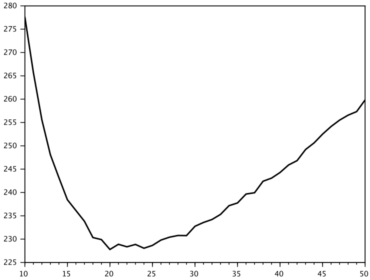


Figure: Averaged number of random variables used in Algorithm (A1) versus the number of slices k with $X_0 = 0$ and $L = 5$. The averaging uses 10 000 simulations.

Example 2: Ornstein-Uhlenbeck process with $b(x) = \alpha x + \beta$, $\alpha = -0.3$, $\beta = 1$ with starting position $X_0 = 0$ and boundary $L = 1$ ensures that γ is a positive function but b remains unbounded. We replace the original drift term by its modified version:

$$b_\rho(x) = \begin{cases} -\alpha x + \beta & \text{if } -\rho \leq x \leq L, \\ \alpha\rho + \beta - \alpha(x+\rho)e^{x+\rho} & \text{if } x < -\rho. \end{cases}$$

The modified γ satisfies $\gamma_\rho(x) = \gamma(x)$ for $x \in [-\rho, L]$ and

$$\gamma_\rho(x) = \frac{1}{2}(\alpha\rho + \beta - \alpha(x+\rho)e^{x+\rho})^2 - \frac{\alpha}{2}(1 + x + \rho)e^{x+\rho} \quad \text{for } x < -\rho.$$

The function γ_ρ is now positive on the whole interval $]-\infty, L]$ and admits the following upper-bound:

$$\kappa = \frac{1}{2}\left(\alpha\rho + \beta + \frac{\alpha}{e}\right)^2 + \frac{\alpha}{2e^2}.$$

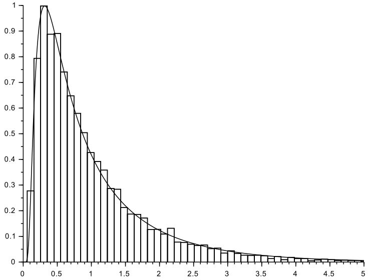
We can therefore apply Algorithm (A1) in order to simulate the approximated first-passage time τ_L^ρ.

We define $\beta(x) = \int_0^x b(y)dy$ and $p(x) = \int_0^x e^{-\beta(y)}\,dy$

Proposition

We assume that $\lim_{x\to-\infty} p(x) = -\infty$. Then τ_L^ρ converges in distribution towards τ_L as $\rho \to \infty$. Moreover

$$d(\tau_L, \tau_L^\rho) := \sup\left\{|F_{\tau_L^\rho}(t) - F_{\tau_L}(t)| \ : \ t \in \mathbb{R}_+\right\} = \mathcal{O}\Big(-p(-\rho)\Big) \text{ as } \rho \to \infty.$$



Histogram of the hitting time distribution for 10 000 simulations, $L = 1$, $X_0 = 0$ for $\rho = 5$ using Algorithm (A1) with modified drift.

Example 3: $dX_t = -\arctan(X_t)\,dt + dB_t, \quad t \geq 0, \quad X_0 = 0, \quad L = 1.$

$\gamma(x) = (\arctan(x)^2 - 1/(1+x^2))/2$ satisfies $-m = -1/2 \leq \gamma(x) \leq \pi^2/8$ and the first-passage time is almost surely finite.

Step 1: *Simulate* T*: distr. of* $(L-x)^2/G^2$ *given* $(L-x)^2/G^2 \leq t_0$.

Step 2: *Simulate a 3-d Bessel process* (R_t) *on* $[0, T]$ *with* $R_T = L - x$.

$$D^m_{R,T} := \Big\{(t, v) \in [0, T] \times \mathbb{R}_+ : \ v \leq \gamma(L - R_t) + \frac{mt_0}{T}\Big\}.$$

Step 3: *Simulate a PPP* N *on* $[0, T] \times \mathbb{R}_+$*, independent of* R*, with Lebesgue intensity meas.*

Step 4: *If* $N(D^m_{R,T}) = 0$ *then set* $Y = T$ *otherwise go to Step 1.*

Theorem for Algorithm (A3)

The outcome Y has the same distribution as τ_L given $\tau_L \leq t_0$.

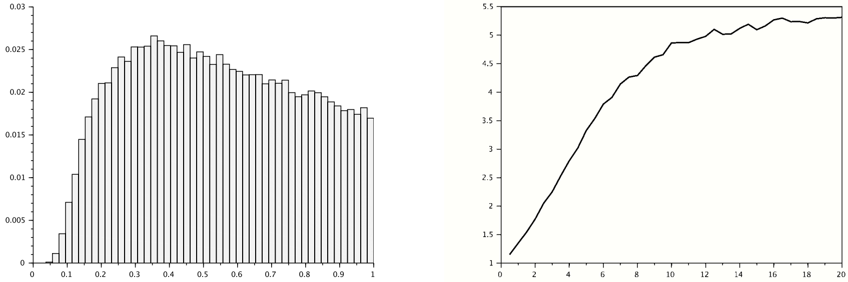


Figure: Histogram of the hitting time distribution using Algorithm (A3) for $t_0 = 1$ and 100 000 simulations (left) and averaged number of iterations in Algorithm (A3) versus t_0 (right) for $X_0 = 0$, $L = 1$ and 10 000 simulations.

To sum up...

Condition on γ	r.v. simulated	Algorithm
$0 \le \gamma(x) \le \kappa$	τ_L	(A1) or (A2)
$0 < \gamma_0 \le \gamma(x) \le \kappa$	τ_L	$(\mathrm{A1})_{\mathrm{shift}}$ or $(\mathrm{A2})_{\mathrm{shift}}$
$-m \le \gamma(x) \le \kappa$	τ_L given $\tau_L \le t_0$	(A3)
$0 \le \gamma(x)$	τ_L^ρ (approx.)	$(\mathrm{A1})^\rho$ or $(\mathrm{A2})^\rho$

To sum up...

Condition on γ	r.v. simulated	Algorithm
$0 \leq \gamma(x) \leq \kappa$	τ_L	(A1) or (A2)
$0 < \gamma_0 \leq \gamma(x) \leq \kappa$	τ_L	$(\mathrm{A1})_{\mathrm{shift}}$ or $(\mathrm{A2})_{\mathrm{shift}}$
$-m \leq \gamma(x) \leq \kappa$	τ_L given $\tau_L \leq t_0$	(A3)
$0 \leq \gamma(x)$	τ_L^ρ (approx.)	$(\mathrm{A1})^\rho$ or $(\mathrm{A2})^\rho$

Work in progress and open questions:

- Exact simulation for unbounded γ, for time-inhomogeneous diffusions.
- Bound of the number of r.v. for general functions γ.
- Exit problem from an interval for one-dimensional diffusions.
- Exit time from a domain in $\mathbb{R}^d$ with $d \geq 2$.