Workshop on Particle and Monte Carlo Methods Barcelona, July 24–25, 2004

Particle methods for the simulation of rare events

> François Le Gland IRISA / INRIA Rennes

legland@irisa.fr
http://www.irisa.fr/aspi/

joint work with Frédéric Cérou (IRISA / INRIA Rennes) Pierre Del Moral (Laboratoire de Statistique et Probabilités, Toulouse) and Pascal Lezaud (Centre d'Études de la Navigation Aérienne, Toulouse)

Plan

- examples of rare events
- multilevel Feynman–Kac distributions
- interacting particle system approximations
- combining importance splitting and importance sampling
- extinction of particle system

example in air-traffic management (ATM) studied in the HYBRIDGE european project (IST programme) partners : NLR, CENA, etc.

- two aircrafts flying over the same area, at the same flight level
- flight plan allows sufficient separation distance between aircrafts
- random perturbations, mainly due to wind, makes actual separation distance smaller than planned separation distance
- risk becomes nonzero, but remains very small
- objective is to evaluate whether flight plan design can be relaxed, so as to increase traffic capacity, without compromising safety

two possible measures of risk

- conflict risk : probability that separation distance gets smaller than 5 nautical miles, roughly 9260 meters
- collision risk : probability that separation distance gets smaller than physical size of aircraft, roughly 100 meters

example in telecommunication networks

- buffer at a station, with service rate much larger than customer arrival rate
- empty buffer is a recurrent event
- large enough buffer size so that overflow, resulting in packet loss, is a rare event
- objective is to evaluate the probability that a buffer overflow occurs, before the buffer empties again

Plan

- examples of rare events
- multilevel Feynman–Kac distributions
- interacting particle system approximations
- combining importance splitting and importance sampling
- extinction of particle system

continuous-time strong Markov process $\{X_t, t \ge 0\}$, with values in metric state space S, and càdlàg trajectories for some closed critical region $B \subset S$, let

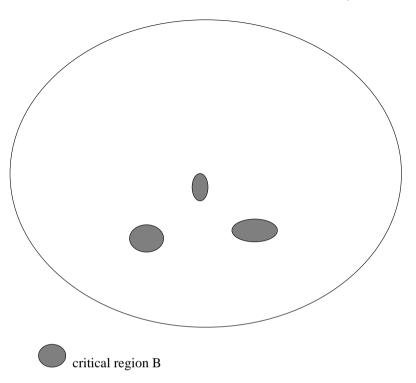
 $T_B = \inf\{t \ge 0 : X_t \in B\}$

with T a finite deterministic time, or an a.s.-finite stopping time objective is to compute probabilities related with rare (but critical) event

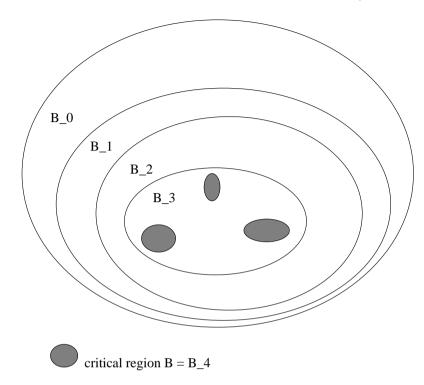
 $\mathbb{P}[T_B \leq T]$ and $\mathbb{E}[f(X_t, 0 \leq t \leq T_B) \mid T_B \leq T]$

i.e.

- probability of the rare event $(T_B \leq T)$
- and probability distribution of the rare trajectories



in practice, none of the simulated trajectories will ever hit the critical region, hence naive Monte Carlo method fails



importance splitting idea is to select trajectories that approach the critical region

introducing an embedded sequence of closed regions

 $B = B_n \subset \cdots \subset B_1 \subset B_0 = S$

with corresponding hitting times

$$T_k = \inf\{t \ge 0 \, : \, X_t \in B_k\}$$

it holds

$$0 = T_0 \le T_1 \le \dots \le T_n = T_B$$

clearly

$$\mathbb{P}[T_B \le T] = \mathbb{P}[T_n \le T] = \prod_{k=0}^n \mathbb{P}[T_k \le T \mid T_{k-1} \le T]$$

but these transition probabilities, from one level to the next, are usually unknown

modelling : discrete-time Markov chain $\{X_k, k = 1 \cdots n\}$ induced by discrete-time events from continuous-time Markov process $\{X_t, t \ge 0\}$

$$\mathfrak{X}_k = (X_t, T_{k-1} \wedge T \le t \le T_k \wedge T)$$

with values in

$$E = \bigcup_{t' \le t''} \mathbb{D}([t', t''], S)$$

for any $e = (x_t, t' \leq t \leq t'') \in E$, let

$$g_k(e) = 1_{(\pi(e) \in B_k)}$$
 where $\pi(e) = x_{t''}$

clearly

$$(T_k \le T)$$
 iff $(X_{T_k \land T} \in B_k)$ iff $(g_k(\mathfrak{X}_k) = 1)$

hence

$$1_{(T_k \le T)} = g_k(\mathfrak{X}_k) = \prod_{p=1}^k g_p(\mathfrak{X}_p)$$

interpretation of rare event probabilities in terms of Feynman-Kac distributions

$$\gamma_k(f) = \mathbb{E}[f(\mathfrak{X}_k) \prod_{p=1}^k g_p(\mathfrak{X}_p)] = \mathbb{E}[f(X_t, T_{k-1} \le t \le T_k) \ 1_{(T_k \le T)}]$$

in particular for $f \equiv 1$

$$\gamma_k(1) = \mathbb{E}\left[\prod_{p=1}^k g_p(\mathfrak{X}_p)\right] = \mathbb{P}[T_k \le T]$$

hence

$$\eta_k(f) = \frac{\gamma_k(f)}{\gamma_k(1)} = \mathbb{E}[f(X_t, T_{k-1} \le t \le T_k) \mid T_k \le T]$$

and in particular

$$\eta_k \circ \pi^{-1}(\phi) = \eta_k(\phi \circ \pi) = \mathbb{E}[\phi(X_{T_k}) \mid T_k \le T]$$

similarly, introducing

$$\gamma_{k|k-1}(f) = \mathbb{E}[f(\mathfrak{X}_k) \prod_{p=1}^{k-1} g_p(\mathfrak{X}_p)] = \mathbb{E}[f(X_t, T_{k-1} \le t \le T_k \land T) \ \mathbf{1}_{(T_{k-1} \le T)}]$$

it holds

$$\eta_{k|k-1}(f) = \frac{\gamma_{k|k-1}(f)}{\gamma_{k|k-1}(1)} = \mathbb{E}[f(X_t, T_{k-1} \le t \le T_k \land T) \mid T_{k-1} \le T]$$

in particular for $f \equiv g_k$

$$\eta_{k|k-1}(g_k) = \frac{\gamma_k(1)}{\gamma_{k-1}(1)} = \mathbb{P}[T_k \le T \mid T_{k-1} \le T]$$

hence

$$\mathbb{P}[T_k \le T] = \gamma_k(1) = \prod_{p=1}^k \eta_{p|p-1}(g_p) = \prod_{p=1}^k \mathbb{P}[T_p \le T \mid T_{p-1} \le T]$$

more generally, introducing path–space Markov chain $\{\chi_{1:k}, k = 1 \cdots n\}$ with $\chi_{1:k} = (\chi_1, \cdots, \chi_k)$, and selection function

$$h_p(e_1, \cdots, e_p) = g_p(e_p)$$
 for any $(e_1, \cdots, e_p) \in E^p$

yields

$$\gamma_{1:k}(f) = \mathbb{E}[f(\mathfrak{X}_{1:k}) \prod_{p=1}^{k} h_p(\mathfrak{X}_{1:p})] = \mathbb{E}[f(\mathfrak{X}_1, \cdots, \mathfrak{X}_k) \prod_{p=1}^{k} g_p(\mathfrak{X}_p)]$$
$$= \mathbb{E}[f(\mathfrak{X}_t, 0 \le t \le T_k) \mathbf{1}_{(T_k} \le T)]$$

and

$$\eta_{1:k}(f) = \frac{\gamma_{1:k}(f)}{\gamma_{1:k}(1)} = \mathbb{E}[f(X_t, 0 \le t \le T_k) \mid T_k \le T]$$

interacting particle methods will provide numerical approximation for rare event probabilities

 $\gamma_n(1) = \mathbb{P}[T_B \leq T]$ and $\eta_{1:n}(f) = \mathbb{E}[f(X_t, 0 \leq t \leq T_B) \mid T_B \leq T]$

and for transition probabilities, from one level to the next

 $\eta_{k|k-1}(g_k) = \mathbb{P}[T_k \le T \mid T_{k-1} \le T]$

with many estimates and asymptotic results as the number of particles goes to infinity, see Section 12.2 in Del Moral, *Feynman–Kac Formulae* (2004)

notice that the selection functions can take the zero value

Plan

- examples of rare events
- multilevel Feynman–Kac distributions
- interacting particle system approximations
- combining importance splitting and importance sampling
- extinction of particle system

Feynman–Kac flow

$$\eta_{k-1} \xrightarrow{\text{prediction}} \eta_{k|k-1} = \eta_{k-1} Q_k \xrightarrow{\text{update}} \eta_k = g_k \cdot \eta_{k|k-1}$$

where \cdot denotes projective product, i.e.

$$\eta_k = g_k \cdot \eta_{k|k-1} = \frac{g_k \eta_{k|k-1}}{\eta_{k|k-1}(g_k)}$$

particle approximation of the form

$$\eta_k \approx \eta_k^N = \sum_{i=1}^N w_k^i \, \delta_{\xi_k^i} \qquad \text{where} \qquad \xi_k^i = (X_t^i \,, \, T_{k-1}^i \wedge T^i \leq t \leq T_k^i \wedge T^i)$$

such that

$$\eta_{k-1}^N \longrightarrow \eta_{k|k-1}^N = S^N(\eta_{k-1}^N Q_k) \longrightarrow \eta_k^N = g_k \cdot \eta_{k|k-1}^N$$

basic (bootstrap) algorithm

• selection of particles with nonzero weight : independently for $i = 1 \cdots N$

 $au_{k-1}^i \sim (w_{k-1}^1, \cdots, w_{k-1}^N)$ with values in index set $\{1, \cdots, N\}$

and

$$\widehat{\xi}_{k-1}^{i} = \xi_{k-1}^{\tau_{k-1}^{i}}$$
 and $\widehat{T}_{k-1}^{i} = T_{k-1}^{\tau_{k-1}^{i}}$

• mutation : independently for $i = 1 \cdots N$

$$\xi_k^i = (X_t^i, \, \widehat{T}_{k-1}^i \le t \le T_k^i \wedge T^i)$$

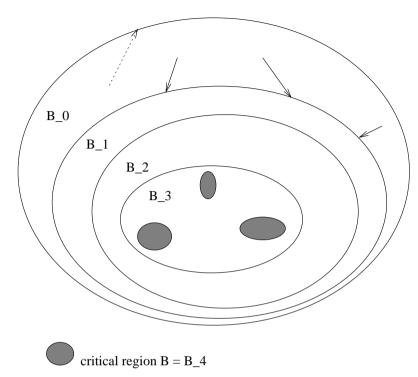
follows the continuous-time Markov model, starting from $\pi(\hat{\xi}_{k-1}^i)$

$$T_k^i = \inf\{t \ge \widehat{T}_{k-1}^i : X_t^i \in B_k\}$$

• weighting according to success to reach next level : for $i = 1 \cdots N$

$$w_k^i \propto g_k(\xi_k^i) = \mathbb{1}_{\left(X_{T_k^i \wedge T^i}^i \in B_k\right)}$$

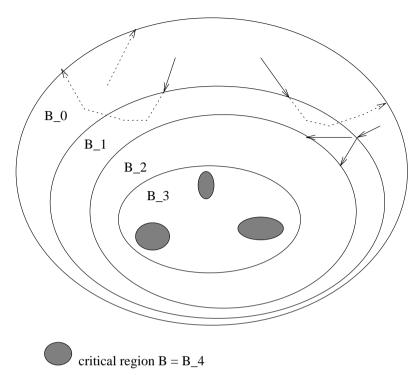
(interacting particle system approximations : 3)



between two levels, particles explore the state space

by mimicking the evolution of the continuous-time Markov process $\{X_t, t \ge 0\}$

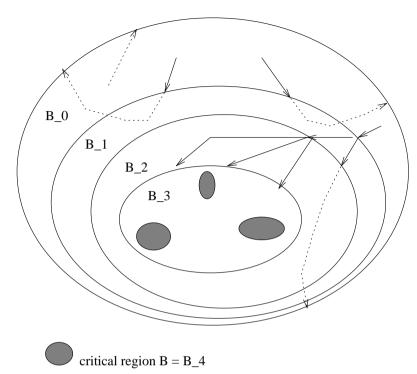
(interacting particle system approximations : 4)



between two levels, particles explore the state space

by mimicking the evolution of the continuous-time Markov process $\{X_t, t \ge 0\}$

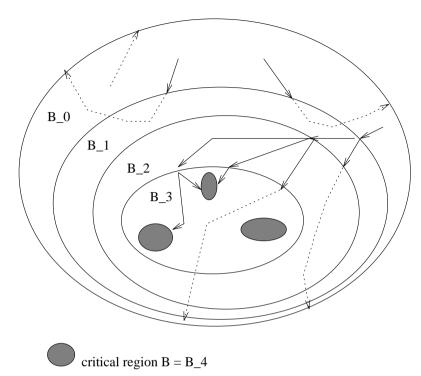
trajectories that succeed to reach the next level before time T are selected, other trajectoires are terminated



between two levels, particles explore the state space

by mimicking the evolution of the continuous-time Markov process $\{X_t, t \ge 0\}$

trajectories that succeed to reach the next level before time T are selected, other trajectoires are terminated



between two levels, particles explore the state space

by mimicking the evolution of the continuous-time Markov process $\{X_t, t \ge 0\}$

trajectories that succeed to reach the next level before time T are selected, other trajectoires are terminated

it could happen that all trajectories fail to reach the next level before time T !

(interacting particle system approximations : 7)

particle approximation of transition probabilities

$$\eta_{k|k-1} \approx \eta_{k|k-1}^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_k^i}$$

and in particular for the test-function g_k

$$\mathbb{P}[T_k \le T \mid T_{k-1} \le T] = \eta_{k|k-1}(g_k) \approx \eta_{k|k-1}^N(g_k) = \frac{1}{N} \sum_{i=1}^N g_k(\xi_k^i) = \frac{|I_k^N|}{N}$$

where

$$I_k^N = \{i = 1 \cdots N : g_k(\xi_k^i) = 1\} = \{i = 1 \cdots N : T_k^i \le T^i\}$$

approximation interpreted as the fraction of trajectories that succeed to reach the next level before time T

(interacting particle system approximations : 8)

particle approximation of rare event probability

$$\gamma_n = \eta_n \prod_{k=1}^n \eta_{k|k-1}(g_k) \approx \gamma_n^N = \eta_n^N \prod_{k=1}^n \eta_{k|k-1}^N(g_k)$$

and in particular

$$\mathbb{P}[T_B \le T] = \gamma_n(1) \approx \gamma_n^N(1) = \prod_{k=1}^n \eta_{k|k-1}^N(g_k) = \prod_{k=1}^n \frac{|I_k^N|}{N}$$

central limit theorem : as $N \uparrow \infty$

$$\sqrt{N} \left(\prod_{k=1}^{n} \frac{|I_k^N|}{N} - \mathbb{P}[T_B \le T]\right) \Longrightarrow \mathcal{N}(0, \sigma_n^2)$$

provided the particle system does not die !

Plan

- examples of rare events
- multilevel Feynman–Kac distributions
- interacting particle system approximations
- combining importance splitting and importance sampling
- extinction of particle system

change of probability measure (e.g. using the Girsanov theorem)

$$\frac{d\mathbb{P}'}{d\mathbb{P}}\Big|_{\mathcal{F}_{T_k}\wedge T} = \prod_{p=1}^k r_p(\mathfrak{X}_p)$$

where

$$\mathcal{F}_t = \sigma(X_s \,,\, 0 \le s \le t)$$

such that, under new probability measure \mathbb{P}' , the event $(T_B \leq T)$ is much less rare

$$\gamma_k(f) = \mathbb{E}[f(\mathfrak{X}_k) \prod_{p=1}^k g_p(\mathfrak{X}_p)] = \mathbb{E}'[f(\mathfrak{X}_k) \prod_{p=1}^k \frac{g_p(\mathfrak{X}_p)}{r_p(\mathfrak{X}_p)}]$$

another Feynman-Kac formula, with

- another continuous-time Markov process, which hits the critical region B with higher probability
- modified selection functions $\frac{g_p}{r_p}$, which penalizes trajectories of the new continuous-time Markov process that do not look like typical trajectories of the original continuous-time Markov process

combined importance splitting / importance sampling algorithm

• selection of particles with higher weights : independently for $i = 1 \cdots N$

 $\tau_{k-1}^i \sim (w_{k-1}^1 \cdots w_{k-1}^N) \qquad \text{with values in index set } \{1, \cdots, N\}$

and

$$\widehat{\xi}_{k-1}^{i} = \xi_{k-1}^{\tau_{k-1}^{i}}$$
 and $\widehat{T}_{k-1}^{i} = T_{k-1}^{\tau_{k-1}^{i}}$

• mutation : independently for $i = 1 \cdots N$

$$\xi_k^i = (X_t^i, \, \widehat{T}_{k-1}^i \le t \le T_k^i \wedge T^i)$$

follows new continuous-time Markov model, starting from $\pi(\hat{\xi}_{k-1}^i)$

$$T_k^i = \inf\{t \ge \widehat{T}_{k-1}^i : X_t^i \in B_k\}$$

• weighting according to success to reach next level and to similarity with a typical trajectory of original continuous-time Markov model : for $i = 1 \cdots N$

$$w_k^i \propto \frac{g_k(\xi_k^i)}{r_k(\xi_k^i)} = \frac{{}^1(X_{T_k^i \wedge T^i}^i \in B_k)}{r_k(\xi_k^i)}$$

particle approximation of transition probabilities

$$\mathbb{P}[T_k \le T \mid T_{k-1} \le T] \approx \frac{1}{N} \sum_{i=1}^N \frac{g_k(\xi_k^i)}{r_k(\xi_k^i)} = \frac{|I_k^N|}{N} \left[\frac{1}{|I_k^N|} \sum_{i \in I_k^N} \frac{1}{r_k(\xi_k^i)}\right]$$

particle approximation of rare event probability

$$\mathbb{P}[T_B \le T] \approx \prod_{k=1}^n \frac{|I_k^N|}{N} \left[\frac{1}{|I_k^N|} \sum_{i \in I_k^N} \frac{1}{r_k(\xi_k^i)} \right]$$

Plan

- examples of rare events
- multilevel Feynman–Kac distributions
- interacting particle system approximations
- combining importance splitting and importance sampling
- extinction of particle system

lifetime of particle system

$$\tau^{N} = \inf\{k \ge 0 : |I_{k}^{N}| = 0\} = \inf\{k \ge 0 : I_{k}^{N} = \emptyset\}$$

if $\mathbb{P}[T_B \leq T] > 0$, then

$$\mathbb{P}[\tau^N \le n] \le c_n \exp\{-a_n N\}$$

for some positive constants $c_n > 0$ and $a_n > 0$

central limit theorem : as $N\uparrow\infty$

$$\sqrt{N} \left(1_{\left(\tau^{N} > n \right)} \prod_{k=1}^{n} \frac{|I_{k}^{N}|}{N} - \mathbb{P}[T_{B} \leq T] \right) \Longrightarrow \mathcal{N}(0, \sigma_{n}^{2})$$

how to make sure that the particle system never dies ?

if it ever dies, reinitialize it as an N–sample from an arbitrary probability distribution ν on E

particle approximation of the form

$$\eta_k \approx \eta_k^N = \sum_{i=1}^N w_k^i \,\delta_{\xi_k^i} \qquad \text{where} \qquad \xi_k^i = (X_t^i, \, T_{k-1}^i \wedge T^i \leq t \leq T_k^i \wedge T^i)$$

such that

$$\eta_{k-1}^N \longrightarrow \eta_{k|k-1}^N = S^N(\eta_{k-1}^N Q_k) \longrightarrow \eta_k^N = g_k \odot \eta_{k|k-1}^N$$

where

$$g_k \odot \eta_{k|k-1}^N = \begin{cases} \frac{g_k \eta_{k|k-1}^N}{\eta_{k|k-1}^N(g_k)} , & \text{if } \eta_{k|k-1}^N(g_k) > 0\\\\ \nu , & \text{otherwise} \end{cases}$$

algorithm with reinitialization, Del Moral, Jacod and Protter (PTRF, 2001)

• selection of particles with nonzero weight : independently for $i=1\cdots N$

 $au_{k-1}^i \sim (w_{k-1}^1, \cdots, w_{k-1}^N)$ with values in index set $\{1, \cdots, N\}$

and

$$\widehat{\xi}_{k-1}^{i} = \xi_{k-1}^{\tau_{k-1}^{i}}$$
 and $\widehat{T}_{k-1}^{i} = T_{k-1}^{\tau_{k-1}^{i}}$

• mutation : independently for $i = 1 \cdots N$

$$\xi_k^i = (X_t^i, \, \widehat{T}_{k-1}^i \le t \le T_k^i \wedge T^i)$$

follows the continuous-time Markov model, starting from $\pi(\widehat{\xi}_{k-1}^{i})$

$$T_k^i = \inf\{t \ge \widehat{T}_{k-1}^i : X_t^i \in B_k\}$$

• if $|I_k^N| \neq 0$, then weighting according to success to reach next level : for $i = 1 \cdots N$

$$w_k^i \propto g_k(\xi_k^i) = \mathbb{1}_{\left(X_{T_k^i \wedge T^i}^i \in B_k\right)}$$

otherwise, if $|I_k^N| = 0$, the whole particle system $\{\xi_k^i, i = 1 \cdots N\}$ is discarded, and reinitialized as an *N*-sample from an arbitrary probability distribution ν on *E*

another way to make sure that the particle systems never dies

generate a random number N_k^H of particles (a stopping time) such that exactly H trajectories succeed to reach the next level

particle approximation of the form

$$\eta_k \approx \eta_k^H = \sum_{i=1}^{N_k^H} w_k^i \,\delta_{\xi_k^i} \qquad \text{where} \qquad \xi_k^i = (X_t^i \,, \, T_{k-1}^i \wedge T^i \leq t \leq T_k^i \wedge T^i)$$

such that

$$\eta_{k-1}^H \longrightarrow \eta_{k|k-1}^H = S^{N_k^H} (\eta_{k-1}^H Q_k) \longrightarrow \eta_k^H = g_k \cdot \eta_{k|k-1}^H$$

where

$$N_k^H = \inf\{N \ge 0 : |I_k^N| = H\} = \inf\{N \ge 0 : \sum_{i=1}^N g_k(\xi_k^i) = H\}$$

sequential algorithm, Oudjane (PhD, 2000), FG and Oudjane (AAP, 2004)

• selection of particles with nonzero weight : independently for $i = 1 \cdots N_k^H$

 $\tau_{k-1}^i \sim (w_{k-1}^1, \cdots, w_{k-1}^{N_{k-1}^H}) \qquad \text{with values in index set } \{1, \cdots, N_{k-1}^H\}$

and

$$\widehat{\xi}_{k-1}^{i} = \xi_{k-1}^{\tau_{k-1}^{i}}$$
 and $\widehat{T}_{k-1}^{i} = T_{k-1}^{\tau_{k-1}^{i}}$

• mutation : independently for $i = 1 \cdots N_k^H$

$$\xi_k^i = (X_t^i, \, \widehat{T}_{k-1}^i \le t \le T_k^i \wedge T^i)$$

follows the continuous-time Markov model, starting from $\pi(\widehat{\xi}_{k-1}^{i})$

$$T_k^i = \inf\{t \ge \widehat{T}_{k-1}^i : X_t^i \in B_k\}$$

• population size N_k^H is chosen such that H trajectories exactly succeed to reach the next level, i.e.

$$N_k^H = \inf\{N \ge 0 : |I_k^N| = H\} = \inf\{N \ge 0 : \sum_{i=1}^N g_k(\xi_k^i) = H\}$$

• weighting according to success to reach next level : for $i = 1 \cdots N_k^H$

$$w_k^i \propto g_k(\xi_k^i) = \mathbb{1}_{\left(X_{T_k^i \wedge T^i}^i \in B_k\right)}$$

given past history of the particle system, the r.v.'s ξ_k^i are i.i.d. : as $N \uparrow \infty$

$$\frac{1}{N} \sum_{i=1}^{N} g_k(\xi_k^i) \longrightarrow \eta_{k-1}^{N_{k-1}^H} Q_k(g_k) = \eta_{k-1}^{N_{k-1}^H} (Q_k g_k)$$

notice that

supp
$$\eta_{k-1}^{N_{k-1}^H} \subset \{e \in E : \pi(e) \in B_{k-1}\}$$

and the strong Markov property yields

$$Q_k g_k(e) = \mathbb{E}[g_k(\mathfrak{X}_k) \mid \mathfrak{X}_{k-1} = e] = \mathbb{P}[T_k \le T \mid X_{T_{k-1} \land T} = \pi(e)]$$

hence, if

$$\mathbb{P}[T_k \le T \mid X_{T_{k-1} \land T} = x] > 0 \qquad \text{for any } x \in B_{k-1}$$

then $\eta_{k-1}^{N_{k-1}^{H}}(Q_{k} g_{k}) > 0$, and

$$N_k^H = \inf\{N \ge 0 : \sum_{i=1}^N g_k(\xi_k^i) = H\}$$

is an a.s. finite and integrable stopping time

particle approximation of transition probabilities

$$\mathbb{P}[T_k \le T \mid T_{k-1} \le T] = \eta_{k|k-1}(g_k) \approx \eta_{k|k-1}^H(g_k) = \frac{|I_k^{N_k^H}|}{N_k^H} = \frac{H}{N_k^H}$$

approximation interpreted as the fraction of trajectories that succeed to reach the next level before time ${\cal T}$

particle approximation of rare event probability

$$\mathbb{P}[T_B \le T] = \gamma_n(1) \approx \gamma_n^H(1) = \prod_{k=1}^n \eta_{k|k-1}^H(g_k) = \prod_{k=1}^n \frac{H}{N_k^H}$$

central limit theorem : as $H\uparrow\infty$

$$\sqrt{H} \left(\prod_{k=1}^{n} \frac{H}{N_{k}^{H}} - \mathbb{P}[T_{B} \leq T]\right) \Longrightarrow \mathcal{N}(0, s_{n}^{2})$$

alternate normalization : average size of particle systems (computing ressources)

$$N_{1:n}^{H} = \frac{1}{n} \sum_{k=1}^{n} N_{k}^{H}$$

central limit theorem : as $H\uparrow\infty$

$$\frac{N_{1:n}^H}{H} \longrightarrow c_n = \frac{1}{n} \sum_{k=1}^n \frac{1}{\eta_{k|k-1}(g_k)}$$

in probability, hence

$$\sqrt{N_{1:n}^H} \left(\prod_{k=1}^n \frac{H}{N_k^H} - \mathbb{P}[T_B \le T]\right) \Longrightarrow \mathcal{N}(0, c_n \, s_n^2)$$

next steps (future work)

- compare σ_n^2 vs. $c_n s_n^2$
- design intermediate levels so as to minimize asymptotic variance