Sampling per mode simulation for switching diffusions

Pascal Lezaud[1] Jaroslav Krystul[2] François Le Gland[3]

[1]DSNA/DTI/R&D and Institut de Mathématiques de Toulouse

[2]Twente University, Enschede

[3]INRIA Rennes, Bretagne–Atlantique

8th International Conference on Rare Event Simulation, RESIM'10, Cambridge, June 21–22, 2010

This work is partially supported by a grant from the European project iFly

Plan

Motivations

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:

Motivations

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:
  - one state component evolves in $\mathbb{R}^d$,

MOTIVATIONS

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:
  - one state component evolves in $\mathbb{R}^d$,
  - the other state component evolves in a discrete set,

Motivations

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:
  - one state component evolves in $\mathbb{R}^d$,
  - the other state component evolves in a discrete set,
  - and each component may influence the evolution of the other component.

MOTIVATIONS

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:
  - one state component evolves in $\mathbb{R}^d$,
  - the other state component evolves in a discrete set,
  - and each component may influence the evolution of the other component.
- Our motivation is to estimate the probability that the continuous component hits a critical set.

Motivations

- Many complex dynamical multi-agent systems make use of continuous-time strong Markov processes with an hybrid state space:
    - one state component evolves in $\mathbb{R}^d$,
    - the other state component evolves in a discrete set,
    - and each component may influence the evolution of the other component.
- Our motivation is to estimate the probability that the continuous component hits a critical set.
- We use a splitting technique adapted to the context of switching diffusions: the sampling per mode algorithm introduced by Krystul in [Krystul, 2006]

Plan

Switching jump diffusion

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,

Switching jump diffusion

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,
- the continuous component is described as a d-dimensional SDE

$$dX_t = b(X_t, \theta_t)\, dt + \sigma(X_t, \theta_t)\, dB_t,$$

SWITCHING JUMP DIFFUSION

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,
- the continuous component is described as a d-dimensional SDE

$$dX_t = b\left(X_t, \theta_t\right) dt + \sigma\left(X_t, \theta_t\right) dB_t,$$

- and the discrete mode as a pure jump process

$$\mathbb{P}\left(\theta_{t+\Delta t} = j | \theta_t = i, X_t = x\right) = \lambda_{ij}(x)\Delta t + o(\Delta t),\ i \neq j,$$

with jump intensities depending on the continuous component.

Switching jump diffusion

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,
- the continuous component is described as a d-dimensional SDE

$$dX_t = b\left(X_t, \theta_t\right) dt + \sigma\left(X_t, \theta_t\right) dB_t,$$

- and the discrete mode as a pure jump process

$$\mathbb{P}\left(\theta_{t+\Delta t} = j | \theta_t = i, X_t = x\right) = \lambda_{ij}(x)\Delta t + o(\Delta t),\ i \neq j,$$

  with jump intensities depending on the continuous component.
- Z_t starts at $t = 0$ in $D_0 \times \mathbb{M}$ with known initial probability η_0

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,
- the continuous component is described as a d-dimensional SDE

$$dX_t = b\left(X_t, \theta_t\right) dt + \sigma\left(X_t, \theta_t\right) dB_t,$$

- and the discrete mode as a pure jump process

$$\mathbb{P}\left(\theta_{t+\Delta t} = j | \theta_t = i, X_t = x\right) = \lambda_{ij}(x)\Delta t + o(\Delta t),\ i \neq j,$$

  with jump intensities depending on the continuous component.
- Z_t starts at $t = 0$ in $D_0 \times \mathbb{M}$ with known initial probability η_0
- Let $A \subset \mathbb{R}^d$ be a closed critical region in which X_t could enter but with a very small probability.

SWITCHING JUMP DIFFUSION

- Strong Markov process $Z = \{(X_t, \theta_t); t \geq 0\}$ with value in $\mathbb{R}^d \times \mathbb{M}$ with a finite set $\mathbb{M} = \{1, \cdots, M\}$,
- the continuous component is described as a d-dimensional SDE

$$dX_t = b(X_t, \theta_t)\, dt + \sigma(X_t, \theta_t)\, dB_t,$$

- and the discrete mode as a pure jump process

$$\mathbb{P}(\theta_{t+\Delta t} = j | \theta_t = i, X_t = x) = \lambda_{ij}(x)\Delta t + o(\Delta t),\ i \neq j,$$

  with jump intensities depending on the continuous component.
- Z_t starts at $t = 0$ in $D_0 \times \mathbb{M}$ with known initial probability η_0
- Let $A \subset \mathbb{R}^d$ be a closed critical region in which X_t could enter but with a very small probability.
- If T_A denotes the hitting time of A, we would like to estimate $\mathbb{P}(T_A \leq T)$ with T a deterministic or a stopping time.

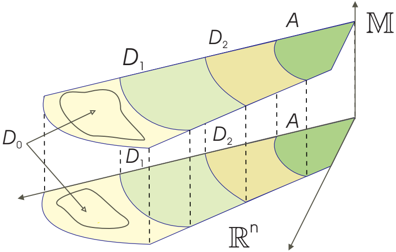
Plan

Splitting technique

- Identify intermediate sets that are (sequentially) visited much more often than the rare target set:

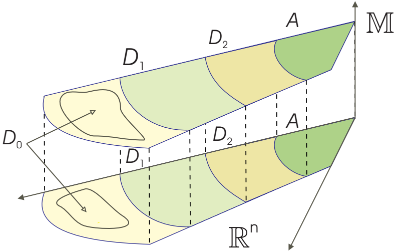
Splitting technique

- Identify intermediate sets that are (sequentially) visited much more often than the rare target set:
- Let $A = D_n \subset \cdots \subset D_1 \subset \mathbb{R}^d, \quad D_0 \cap D_1 = \emptyset$



Splitting technique

- Identify intermediate sets that are (sequentially) visited much more often than the rare target set:
- Let $A = D_n \subset \cdots \subset D_1 \subset \mathbb{R}^d, \quad D_0 \cap D_1 = \emptyset$



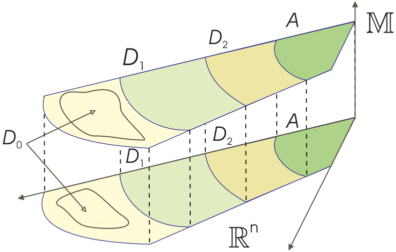
- With $B = A \times \mathbb{M}$ and $B_k = D_k \times \mathbb{M}$, we define for $k = 1, \cdots, n$

$$T_k = \inf\{t \geq 0 : Z_t \in B_k\} = \inf\{t \geq 0 : X_t \in D_k\},$$

which satisfy $0 = T_0 \leq T_1 \leq \cdots \leq T_n = T_B$.

Splitting technique

- Identify intermediate sets that are (sequentially) visited much more often than the rare target set:
- Let $A = D_n \subset \cdots \subset D_1 \subset \mathbb{R}^d, \quad D_0 \cap D_1 = \emptyset$



- With $B = A \times \mathbb{M}$ and $B_k = D_k \times \mathbb{M}$, we define for $k = 1, \cdots, n$

$$T_k = \inf\{t \geq 0 : Z_t \in B_k\} = \inf\{t \geq 0 : X_t \in D_k\},$$

which satisfy $0 = T_0 \leq T_1 \leq \cdots \leq T_n = T_B$.

- Then

$$\mathbb{P}(T_A \leq T) = \mathbb{P}(T_B \leq T) = \prod_{k=1}^{n} \mathbb{P}(T_k \leq T | T_{k-1} \leq T),$$

where conditional probabilities are not very small.

Plan

Some issues

- Splitting technique applies since a switching process is a strong Markov process, but

Some issues

- Splitting technique applies since a switching process is a strong Markov process, but
- this approach fails to produce a reasonable estimate, since each resampling step tends

Some issues

- Splitting technique applies since a switching process is a strong Markov process, but
- this approach fails to produce a reasonable estimate, since each resampling step tends
  - to sample more particles from mode with higher probability,

Some issues

- Splitting technique applies since a switching process is a strong Markov process, but
- this approach fails to produce a reasonable estimate, since each resampling step tends
  - to sample more particles from mode with higher probability,
  - to discard the particles in the ”light” modes.

SOME ISSUES

- Splitting technique applies since a switching process is a strong Markov process, but
- this approach fails to produce a reasonable estimate, since each resampling step tends
  - to sample more particles from mode with higher probability,
  - to discard the particles in the ”light” modes.
- Increasing the number of particles should improve the estimate but only at the cost of increased simulation time,

Some issues

- Splitting technique applies since a switching process is a strong Markov process, but
- this approach fails to produce a reasonable estimate, since each resampling step tends
  - to sample more particles from mode with higher probability,
  - to discard the particles in the ”light” modes.
- Increasing the number of particles should improve the estimate but only at the cost of increased simulation time,
- Idea: keep constant the number of particles in each visited mode at each resampling step,

Plan

Multilevel Feynman-Kac distributions

- To capture the behaviour of Z between each thresholds, we consider the random excursions $\mathcal{Z}_k$ of Z between T_{k-1} and $T_k \wedge T$

$$\mathcal{Z}_k = ((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k \wedge T)\,,$$

Multilevel Feynman-Kac distributions

- To capture the behaviour of Z between each thresholds, we consider the random excursions $\mathcal{Z}_k$ of Z between T_{k-1} and $T_k \wedge T$

$$\mathcal{Z}_k = \left((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k \wedge T\right),$$

- and we introduce the selection functions,

$$g_k(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in B_k\}}, \quad g_k^j(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in D_k \times \{j\}\}}, \quad j \in \mathbb{M},$$

Multilevel Feynman-Kac distributions

- To capture the behaviour of Z between each thresholds, we consider the random excursions $\mathcal{Z}_k$ of Z between T_{k-1} and $T_k \wedge T$
$$\mathcal{Z}_k = \left((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k \wedge T\right),$$
- and we introduce the selection functions,
$$g_k(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in B_k\}}, \quad g_k^j(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in D_k \times \{j\}\}}, \quad j \in \mathbb{M},$$

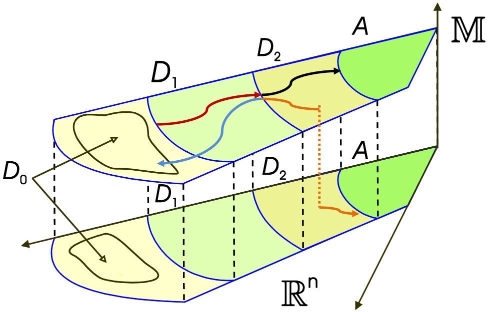
Multilevel Feynman-Kac distributions

- To capture the behaviour of Z between each thresholds, we consider the random excursions $\mathcal{Z}_k$ of Z between T_{k-1} and $T_k \wedge T$

$$\mathcal{Z}_k = ((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k \wedge T)\,,$$

- and we introduce the selection functions,

$$g_k(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in B_k\}}, \quad g_k^j(\mathcal{Z}_k) = 1_{\{Z_{T_k \wedge T} \in D_k \times \{j\}\}}, \quad j \in \mathbb{M},$$



Multilevel Feynman-Kac distributions

- Clearly,

$$1_{\{T_k \leq T\}} = g_k(\mathcal{Z}_k), \quad \text{and} \quad 1_{\{T_k \leq T, \theta_{T_k} = j\}} = g_k^j(\mathcal{Z}_k).$$

- Clearly,

$$1_{\{T_k \leq T\}} = g_k(\mathcal{Z}_k), \text{ and } 1_{\{T_k \leq T, \theta_{T_k} = j\}} = g_k^j(\mathcal{Z}_k).$$

- We can interpret the rare event probability in terms of the Feynman-Kac measures defined by

$$\gamma_k(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_{k-1}(\mathcal{Z}_{k-1})\right] = \mathbb{E}\left[f((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k \wedge T) 1_{\{T_{k-1} \leq T\}}\right]$$

$$\widehat{\gamma}_k(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_k(\mathcal{Z}_k)\right] = \mathbb{E}\left[f((X_t, \theta_t),\ T_{k-1} \leq t \leq T_k) 1_{\{T_k \leq T\}}\right],$$

Multilevel Feynman-Kac distributions

- Clearly,

$$1_{\{T_k \leq T\}} = g_k(\mathcal{Z}_k), \text{ and } 1_{\{T_k \leq T, \theta_{T_k} = j\}} = g_k^j(\mathcal{Z}_k).$$

- We can interpret the rare event probability in terms of the Feynman-Kac measures defined by

$$\gamma_k(f) = \mathbb{E}\left[f(\mathcal{Z}_k)g_{k-1}(\mathcal{Z}_{k-1})\right] = \mathbb{E}\left[f((X_t,\theta_t),\ T_{k-1} \leq t \leq T_k \wedge T)1_{\{T_{k-1} \leq T\}}\right]$$

$$\widehat{\gamma}_k(f) = \mathbb{E}\left[f(\mathcal{Z}_k)g_k(\mathcal{Z}_k)\right] = \mathbb{E}\left[f((X_t,\theta_t),\ T_{k-1} \leq t \leq T_k)1_{\{T_k \leq T\}}\right],$$

- and the corresponding normalized measures defined by

$$\eta_k(f) = \frac{\gamma_k(f)}{\gamma_k(1)} = \mathbb{E}\left[f((X_t,\theta_t), T_{k-1} \leq t \leq T_k \wedge T) | T_{k-1} \leq T\right]$$

$$\widehat{\eta}_k(f) = \frac{\widehat{\gamma}_k(f)}{\widehat{\gamma}_k(1)} = \mathbb{E}\left[f((X_t,\theta_t),\ T_{k-1} \leq t \leq T_k) | T_k \leq T\right].$$

Multilevel Feynman-Kac distributions

- In particular, for $f \equiv 1$

$$\gamma_k(1) = \mathbb{P}(T_{k-1} \leq T), \quad \widehat{\gamma}_k(1) = \mathbb{P}(T_k \leq T),$$

Multilevel Feynman-Kac distributions

- In particular, for $f \equiv 1$

$$\gamma_k(1) = \mathbb{P}(T_{k-1} \leq T), \quad \widehat{\gamma}_k(1) = \mathbb{P}(T_k \leq T),$$

- and for $f = g_k$ or $f = g_k^j$, it holds

$$\eta_k(g_k) = \mathbb{P}[T_k \leq T | T_{k-1} \leq T], \quad \eta_k(g_k^j) = \mathbb{P}[T_k \leq T, \theta_{T_k} = j | T_{k-1} \leq T].$$

Multilevel Feynman-Kac distributions

- In particular, for $f \equiv 1$

$$\gamma_k(1) = \mathbb{P}(T_{k-1} \leq T), \quad \widehat{\gamma}_k(1) = \mathbb{P}(T_k \leq T),$$

- and for $f = g_k$ or $f = g_k^j$, it holds

$$\eta_k(g_k) = \mathbb{P}[T_k \leq T | T_{k-1} \leq T], \quad \eta_k(g_k^j) = \mathbb{P}[T_k \leq T, \theta_{T_k} = j | T_{k-1} \leq T].$$

- We have the key formulas

$$\gamma_k(f) = \eta_k(f) \prod_{p=0}^{k-1} \eta_p(g_p) \quad \text{and} \quad \widehat{\gamma}_k(f) = \widehat{\eta}_k(f) \prod_{p=0}^{k} \eta_p(g_p).$$

Multilevel Feynman-Kac distributions

- In particular, for $f \equiv 1$

$$\gamma_k(1) = \mathbb{P}(T_{k-1} \leq T), \quad \widehat{\gamma}_k(1) = \mathbb{P}(T_k \leq T),$$

- and for $f = g_k$ or $f = g_k^j$, it holds

$$\eta_k(g_k) = \mathbb{P}[T_k \leq T | T_{k-1} \leq T], \quad \eta_k(g_k^j) = \mathbb{P}[T_k \leq T, \theta_{T_k} = j | T_{k-1} \leq T].$$

- We have the key formulas

$$\gamma_k(f) = \eta_k(f) \prod_{p=0}^{k-1} \eta_p(g_p) \quad \text{and} \quad \widehat{\gamma}_k(f) = \widehat{\eta}_k(f) \prod_{p=0}^{k} \eta_p(g_p).$$

- Then, we recover

$$\mathbb{P}(T_n \leq T) = \prod_{p=0}^{n} \eta_p(g_p),$$

Multilevel Feynman-Kac distributions

- In order to keep trace of the discrete mode, we construct for any $j \in \mathbb{M}$

$$\gamma_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_{k-1}^j(\mathcal{Z}_{k-1})\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k \wedge T) 1_{\{T_{k-1} \leq T, \theta_{T_{k-1}} = j\}}\right]$$

$$\widehat{\gamma}_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_k^j(\mathcal{Z}_k)\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k) 1_{\{T_k \leq T, \theta_{T_k} = j\}}\right].$$

Multilevel Feynman-Kac distributions

- In order to keep trace of the discrete mode, we construct for any $j \in \mathbb{M}$

$$\gamma_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k)g_{k-1}^j(\mathcal{Z}_{k-1})\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k \wedge T)1_{\{T_{k-1}\leq T, \theta_{T_{k-1}}=j\}}\right]$$

$$\widehat{\gamma}_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k)g_k^j(\mathcal{Z}_k)\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k)1_{\{T_k \leq T, \theta_{T_k} = j\}}\right].$$

- and the normalized measures

$$\eta_k^j(f) = \frac{\gamma_k^j(f)}{\gamma_k^j(1)} = \mathbb{E}\left[f(Z_t,\ T_{k-1} \leq t \leq T_k \wedge T) | T_{k-1} \leq T, \theta_{T_{k-1}} = j\right],$$

$$\widehat{\eta}_k^j(f) = \frac{\widehat{\gamma}_k^j(f)}{\widehat{\gamma}_k^j(1)} = \mathbb{E}\left[f(Z_t,\ T_{k-1} \leq t \leq T_k) | T_k \leq T, \theta_{T_k} = j\right].$$

MULTILEVEL FEYNMAN-KAC DISTRIBUTIONS

- In order to keep trace of the discrete mode, we construct for any $j \in \mathbb{M}$

$$\gamma_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_{k-1}^j(\mathcal{Z}_{k-1})\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k \wedge T) 1_{\{T_{k-1} \leq T, \theta_{T_{k-1}} = j\}}\right]$$

$$\widehat{\gamma}_k^j(f) = \mathbb{E}\left[f(\mathcal{Z}_k) g_k^j(\mathcal{Z}_k)\right] = \mathbb{E}\left[f(Z_t, T_{k-1} \leq t \leq T_k) 1_{\{T_k \leq T, \theta_{T_k} = j\}}\right].$$

- and the normalized measures

$$\eta_k^j(f) = \frac{\gamma_k^j(f)}{\gamma_k^j(1)} = \mathbb{E}\left[f(Z_t,\ T_{k-1} \leq t \leq T_k \wedge T) | T_{k-1} \leq T, \theta_{T_{k-1}} = j\right],$$

$$\widehat{\eta}_k^j(f) = \frac{\widehat{\gamma}_k^j(f)}{\widehat{\gamma}_k^j(1)} = \mathbb{E}\left[f(Z_t,\ T_{k-1} \leq t \leq T_k) | T_k \leq T, \theta_{T_k} = j\right].$$

- We have the decompositions

$$\widehat{\eta}_k = \sum_{j \in \mathbb{M}} \omega_k^j \widehat{\eta}_k^j, \qquad \eta_{k+1} = \sum_{j \in \mathbb{M}} \omega_k^j \eta_{k+1}^j,$$

where

$$\omega_k^j = \widehat{\eta}_k(g_k^j) = \mathbb{P}(\theta_{T_k} = j | T_k \leq T).$$

Plan

- Using the Markov property of $\mathcal{Z}$ (with Markov kernel $\mathcal{M}_k$), we obtain

$$\gamma_k(f) = \gamma_{k-1}(g_{k-1}\mathcal{M}_k\, f) \text{ and } \gamma_k^j = \gamma_{k-1}(g_{k-1}^j\mathcal{M}_k\, f).$$

- Using the Markov property of $\mathcal{Z}$ (with Markov kernel $\mathcal{M}_k$), we obtain

$$\gamma_k(f) = \gamma_{k-1}(g_{k-1}\mathcal{M}_k\, f) \text{ and } \gamma_k^j = \gamma_{k-1}(g_{k-1}^j\mathcal{M}_k\, f).$$

- and the nonlinear measure-valued transformations

$$\widehat{\eta}_k(f) = \frac{\eta_k(fg_k)}{\eta_k(g_k)} := \Psi_k(\eta_k)(f), \qquad \widehat{\eta}_k^j(f) = \frac{\eta_k(fg_k^j)}{\eta_k(g_k^j)} := \Psi_k^j(\eta_k)(f).$$

Dynamical evolution

- Using the Markov property of $\mathcal{Z}$ (with Markov kernel $\mathcal{M}_k$), we obtain

$$\gamma_k(f) = \gamma_{k-1}(g_{k-1}\mathcal{M}_k\, f) \text{ and } \gamma_k^j = \gamma_{k-1}(g_{k-1}^j\mathcal{M}_k\, f).$$

- and the nonlinear measure-valued transformations

$$\widehat{\eta}_k(f) = \frac{\eta_k(fg_k)}{\eta_k(g_k)} := \Psi_k(\eta_k)(f), \qquad \widehat{\eta}_k^j(f) = \frac{\eta_k(fg_k^j)}{\eta_k(g_k^j)} := \Psi_k^j(\eta_k)(f).$$

- so, the following two separate selection/mutation transitions

$$\eta_k \xrightarrow{\text{selection}} \widehat{\eta}_k := \Psi_k(\eta_k) \xrightarrow{\text{mutation}} \eta_{k+1} = \widehat{\eta}_k\mathcal{M}_{k+1}.$$

Plan

Particle methods

- Particle methods are a kind of stochastic linearisation technique for solving nonlinear equation in measure space.

Particle methods

- Particle methods are a kind of stochastic linearisation technique for solving nonlinear equation in measure space.
- Using two sequences of N particles $\xi = (\xi^1, \cdots, \xi^N)$ and $\widehat{\xi} = (\widehat{\xi}^1, \cdots, \widehat{\xi}^N)$, we approximate the two step transitions

PARTICLE METHODS

- Particle methods are a kind of stochastic linearisation technique for solving nonlinear equation in measure space.
- Using two sequences of N particles $\xi = (\xi^1, \cdots, \xi^N)$ and $\widehat{\xi} = (\widehat{\xi}^1, \cdots, \widehat{\xi}^N)$, we approximate the two step transitions

- Particle methods are a kind of stochastic linearisation technique for solving nonlinear equation in measure space.
- Using two sequences of N particles $\xi = (\xi^1, \cdots, \xi^N)$ and $\widehat{\xi} = (\widehat{\xi}^1, \cdots, \widehat{\xi}^N)$, we approximate the two step transitions

$$\eta_k \xrightarrow{\text{selection}} \widehat{\eta}_k := \Psi_k(\eta_k) \xrightarrow{\text{mutation}} \eta_{k+1} = \widehat{\eta}_k \mathcal{M}_{k+1},$$

Particle methods

- Particle methods are a kind of stochastic linearisation technique for solving nonlinear equation in measure space.
- Using two sequences of N particles $\xi = (\xi^1, \cdots, \xi^N)$ and $\widehat{\xi} = (\widehat{\xi}^1, \cdots, \widehat{\xi}^N)$, we approximate the two step transitions

$$\eta_k \xrightarrow{\text{selection}} \widehat{\eta}_k := \Psi_k(\eta_k) \xrightarrow{\text{mutation}} \eta_{k+1} = \widehat{\eta}_k \mathcal{M}_{k+1},$$

by

$$\eta_k^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_k^i} \xrightarrow{\text{selection}} \widehat{\eta}_k^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\widehat{\xi}_k^i} \xrightarrow{\text{mutation}} \eta_{k+1} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_{k+1}^i}.$$

Plan

Sampling per Mode algorithm

- The main idea of Sampling per mode algorithm consists in maintaining a fixed number of particles in each mode, at each resampling step.

Sampling per Mode algorithm

- The main idea of Sampling per mode algorithm consists in maintaining a fixed number of particles in each mode, at each resampling step.
- So, instead of starting the algorithm with N particles randomly distributed, we draw in each mode j, a fixed number N^j particles and at each resampling step, the same number of particles is sampled for each visited mode.

Sampling per Mode algorithm

- The main idea of Sampling per mode algorithm consists in maintaining a fixed number of particles in each mode, at each resampling step.
- So, instead of starting the algorithm with N particles randomly distributed, we draw in each mode j, a fixed number N^j particles and at each resampling step, the same number of particles is sampled for each visited mode.
- Obviously, the total number of particles can change at each time some mode is not visited, or empty mode is visited afresh.

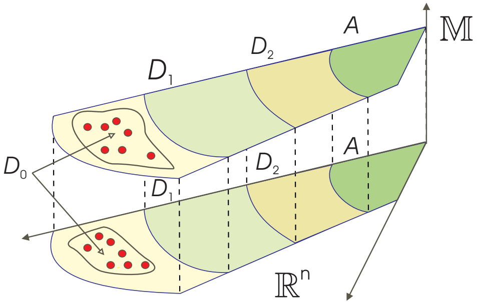
Sampling per Mode algorithm

- The main idea of Sampling per mode algorithm consists in maintaining a fixed number of particles in each mode, at each resampling step.
- So, instead of starting the algorithm with N particles randomly distributed, we draw in each mode j, a fixed number N^j particles and at each resampling step, the same number of particles is sampled for each visited mode.
- Obviously, the total number of particles can change at each time some mode is not visited, or empty mode is visited afresh.
- Let $\widehat{N}_k$ and N_k denote the total numbers of particles $\widehat{\xi}_k$ and ξ_k, and $\omega_k^{j,N}$ the weights associated with the modes, we have the evolution scheme

$$(N_k, (\omega_{k-1}^{j,N})_{j \in J_{k-1}}, \xi_k) \to (\widehat{N}_k, (\omega_k^j)_{j \in J_k}, \widehat{\xi}_k) \to (N_{k+1}, (\omega_k^{j,N})_{j \in J_k}, \xi_{k+1})$$

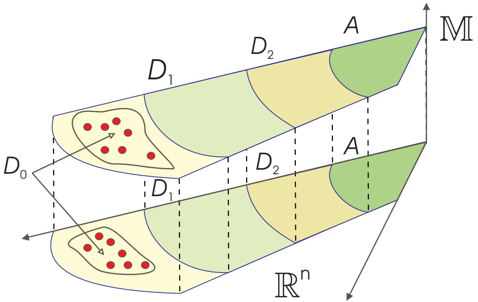
where J_k denotes the set of non empty modes at step k

- In each mode j, we sample N^j particles $\xi_0^\kappa = \widehat{\xi}_0^\kappa = (0, (X_0^\kappa, j)) \sim \eta_0^j$.



Initialization

- In each mode j, we sample N^j particles $\xi_0^\kappa = \widehat{\xi}_0^\kappa = (0, (X_0^\kappa, j)) \sim \eta_0^j$.



- Let $\omega_0^j = \mathbb{P}(\theta_0 = j)$, then η_0^N and $\widehat{\eta}_0^N$ are given by

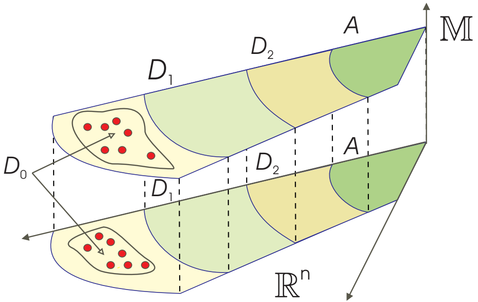
$$\eta_0^N = \sum_{j\in\mathbb{M}} \omega_0^j \eta_0^{j,N}, \qquad \widehat{\eta}_0^N = \sum_{j\in\mathbb{M}} \omega_0^j \widehat{\eta}_0^{j,N},$$

with

$$\eta_0^{j,N} = \frac{1}{N^j} \sum_{\kappa\in J_0^j} \delta_{\xi_0^\kappa}, \qquad \widehat{\eta}_0^{j,N} = \frac{1}{N^j} \sum_{\kappa\in J_0^j} \delta_{\widehat{\xi}_0^\kappa}.$$

Initialization

- In each mode j, we sample N^j particles $\xi_0^\kappa = \widehat{\xi}_0^\kappa = (0, (X_0^\kappa, j)) \sim \eta_0^j$.



- Let $\omega_0^j = \mathbb{P}(\theta_0 = j)$, then η_0^N and $\widehat{\eta}_0^N$ are given by

$$\eta_0^N = \sum_{j\in\mathbb{M}} \omega_0^j \eta_0^{j,N}, \qquad \widehat{\eta}_0^N = \sum_{j\in\mathbb{M}} \omega_0^j \widehat{\eta}_0^{j,N},$$

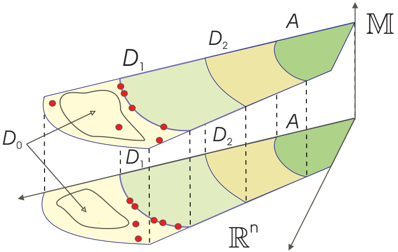
with

$$\eta_0^{j,N} = \frac{1}{N^j} \sum_{\kappa\in J_0^j} \delta_{\xi_0^\kappa}, \qquad \widehat{\eta}_0^{j,N} = \frac{1}{N^j} \sum_{\kappa\in J_0^j} \delta_{\widehat{\xi}_0^\kappa}.$$

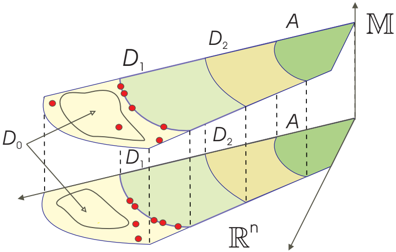
- Here J_0^j is the set of the indices of the particles in the mode j.

Mutation $(\widehat{N}_k, \omega_k^N, \widehat{\xi}_k) \to (N_{k+1}, \omega_k^N, \xi_{k+1})$

- If $\widehat{N}_k = 0$ the particle system dies, otherwise independently of each other, each particle $\widehat{\xi}_k^\kappa$ evolves randomly according to the Markov transition $\mathcal{M}_{k+1}$



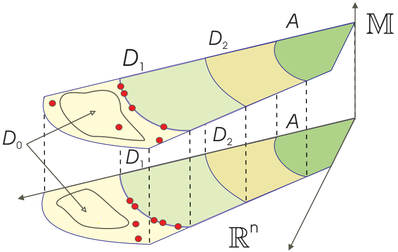
- If $\widehat{N}_k = 0$ the particle system dies, otherwise independently of each other, each particle $\widehat{\xi}_k^\kappa$ evolves randomly according to the Markov transition $\mathcal{M}_{k+1}$



- Neither the total number of particles nor the weight of each particle change ($N_{k+1} = \widehat{N}_k$).

Mutation $(\widehat{N}_k, \omega_k^N, \widehat{\xi}_k) \to (N_{k+1}, \omega_k^N, \xi_{k+1})$

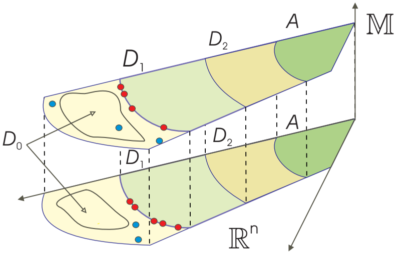
- If $\widehat{N}_k = 0$ the particle system dies, otherwise independently of each other, each particle $\widehat{\xi}_k^\kappa$ evolves randomly according to the Markov transition $\mathcal{M}_{k+1}$



- Neither the total number of particles nor the weight of each particle change ($N_{k+1} = \widehat{N}_k$).
- So $\eta_{k+1}^N = \sum_{j \in J_k} \omega_k^{j,N} \eta_{k+1}^{j,N}$, with $\eta_{k+1}^{j,N} = \frac{1}{N^j} \sum_{\kappa \in J_k^j} \delta_{\xi_{k+1}^\kappa}$, where J_k^j is the set of the labels of the particles in mode $j \in J_k$.

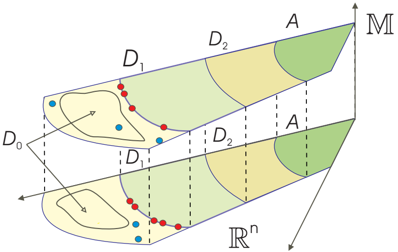
SELECTION/RESAMPLING $(N_{k+1}, \omega_k^N, \xi_{k+1}) \rightarrow (\widehat{N}_{k+1}, \omega_{k+1}^N, \widehat{\xi}_{k+1})$

- Select only the particles ξ_{k+1}^{κ} having reached the desired set B_{k+1};



SELECTION/RESAMPLING $(N_{k+1}, \omega_k^N, \xi_{k+1}) \to (\widehat{N}_{k+1}, \omega_{k+1}^N, \widehat{\xi}_{k+1})$

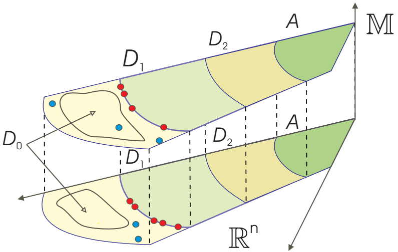
- Select only the particles ξ_{k+1}^{κ} having reached the desired set B_{k+1};



- Let I_{k+1}^N denote the set of (indices of) good particles; if $I_{k+1}^N = \emptyset$ the algorithm is stopped. Otherwise,

SELECTION/RESAMPLING $(N_{k+1}, \omega_k^N, \xi_{k+1}) \to (\widehat{N}_{k+1}, \omega_{k+1}^N, \widehat{\xi}_{k+1})$

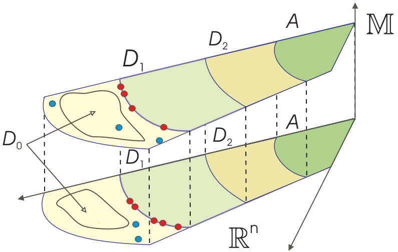
- Select only the particles ξ_{k+1}^{κ} having reached the desired set B_{k+1};



- Let I_{k+1}^N denote the set of (indices of) good particles; if $I_{k+1}^N = \emptyset$ the algorithm is stopped. Otherwise,
- for each non empty mode j, resample N^j particles according to $\Psi_{k+1}^j(\eta_{k+1}^N)$, and set

SELECTION/RESAMPLING $(N_{k+1}, \omega_k^N, \xi_{k+1}) \to (\widehat{N}_{k+1}, \omega_{k+1}^N, \widehat{\xi}_{k+1})$

- Select only the particles ξ_{k+1}^{κ} having reached the desired set B_{k+1};

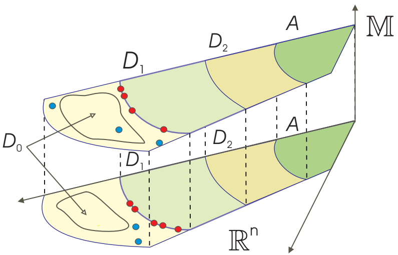


- Let I_{k+1}^N denote the set of (indices of) good particles; if $I_{k+1}^N = \emptyset$ the algorithm is stopped. Otherwise,
- for each non empty mode j, resample N^j particles according to $\Psi_{k+1}^j(\eta_{k+1}^N)$, and set
- $\widehat{\eta}_{k+1}^N = \sum_{j \in J_{k+1}} \omega_{k+1}^{j,N} \widehat{\eta}_{k+1}^{j,N}$, with $\widehat{\eta}_{k+1}^{j,N} = \frac{1}{N^j} \sum_{\kappa \in \widehat{J}_{k+1}^j} \delta_{\widehat{\xi}_{k+1}^{\kappa}}$ and

$$\omega_{k+1}^{j,N} = \widehat{\eta}_{k+1}^N(g_{k+1}^j) = \frac{\eta_{k+1}^N(g_{k+1}^j)}{\eta_{k+1}^N(g_{k+1})}.$$

SELECTION/RESAMPLING $(N_{k+1}, \omega_k^N, \xi_{k+1}) \to (\widehat{N}_{k+1}, \omega_{k+1}^N, \widehat{\xi}_{k+1})$

- Select only the particles ξ_{k+1}^κ having reached the desired set B_{k+1};

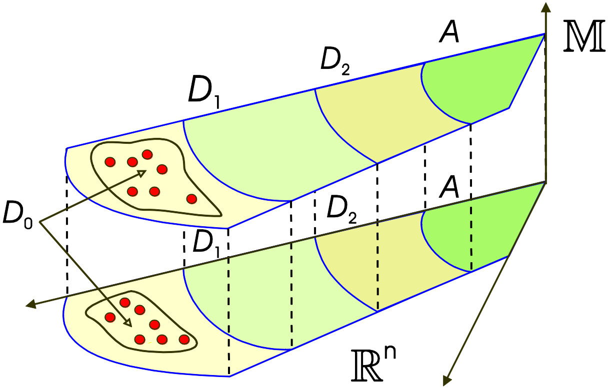


- Let I_{k+1}^N denote the set of (indices of) good particles; if $I_{k+1}^N = \emptyset$ the algorithm is stopped. Otherwise,
- for each non empty mode j, resample N^j particles according to $\Psi_{k+1}^j(\eta_{k+1}^N)$, and set
- $\widehat{\eta}_{k+1}^N = \sum_{j \in J_{k+1}} \omega_{k+1}^{j,N} \widehat{\eta}_{k+1}^{j,N}$, with $\widehat{\eta}_{k+1}^{j,N} = \frac{1}{N^j} \sum_{\kappa \in \widehat{J}_{k+1}^j} \delta_{\widehat{\xi}_{k+1}^\kappa}$ and

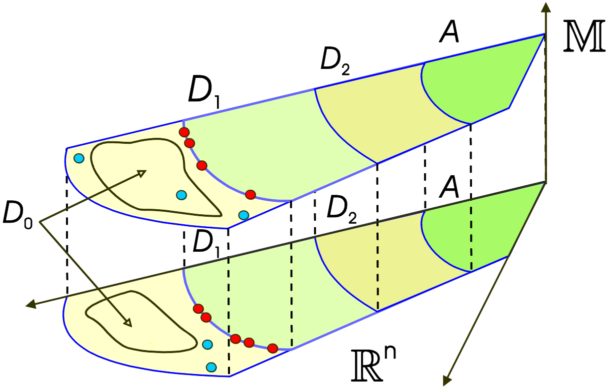
$$\omega_{k+1}^{j,N} = \widehat{\eta}_{k+1}^N(g_{k+1}^j) = \frac{\eta_{k+1}^N(g_{k+1}^j)}{\eta_{k+1}^N(g_{k+1})}.$$

- The total number of particles is $\widehat{N}_{k+1} = \sum_{j \in J_{k+1}} N^j$.

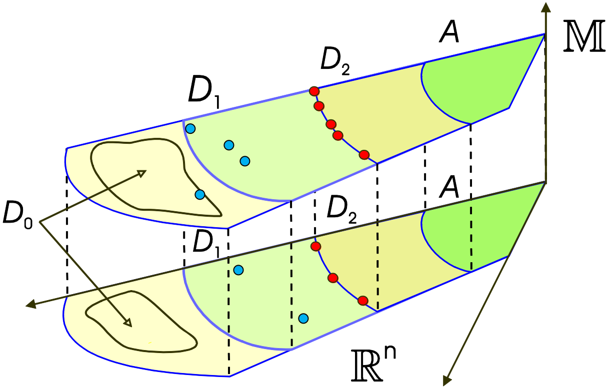
Sampling per Mode algorithm: Recapitulation



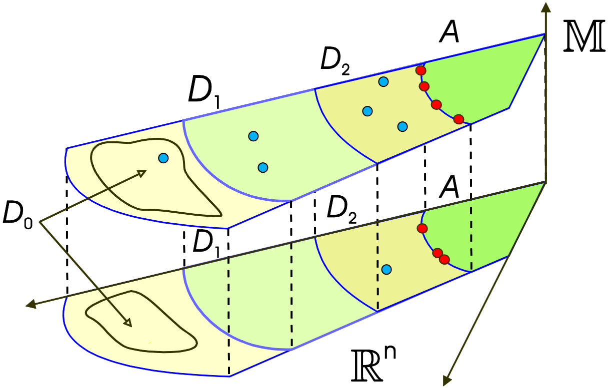
Sampling per Mode algorithm: Recapitulation



Sampling per Mode algorithm: Recapitulation



Sampling per Mode algorithm: Recapitulation



Plan

Asymptotic Behaviour

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,

Asymptotic Behaviour

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,
- and for the central limit theorem, we used a CLT for triangular arrays developed in [Le Gland & Oudjane, 2006]

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,
- and for the central limit theorem, we used a CLT for triangular arrays developed in [Le Gland & Oudjane, 2006]
- Before the statement of the two theorems, we need some notations:

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,
- and for the central limit theorem, we used a CLT for triangular arrays developed in [Le Gland & Oudjane, 2006]
- Before the statement of the two theorems, we need some notations:
  - $N_{\text{inf}} = \inf_{j \in \mathbb{M}} N^j$

Asymptotic Behaviour

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,
- and for the central limit theorem, we used a CLT for triangular arrays developed in [Le Gland & Oudjane, 2006]
- Before the statement of the two theorems, we need some notations:
  - $N_{\text{inf}} = \inf_{j \in \mathbb{M}} N^j$
  - let $N \to \infty$ in such a way that each $\rho_j := N^j/N$ are "preserved"

- Now, we are addressing the asymptotic behaviour of our estimator as $N \to \infty$.
- To obtain a law of large numbers, we followed the [Del Moral 2004]'s approach based on a martingale decomposition,
- and for the central limit theorem, we used a CLT for triangular arrays developed in [Le Gland & Oudjane, 2006]
- Before the statement of the two theorems, we need some notations:
  - $N_{\text{inf}} = \inf_{j \in \mathbb{M}} N^j$
  - let $N \to \infty$ in such a way that each $\rho_j := N^j / N$ are "preserved"
  - this implies that $N_{\text{inf}} \to \infty$.

Plan

Main Theorems: Law of Large Numbers

Theorem (Law of Large Numbers)

For any $n \geq 0$ and any bounded function f, we have

$$\mathbb{E}(\gamma_n^N(f) 1_{\{N_n > 0\}}) = \gamma_n(f),$$

Theorem (Law of Large Numbers)

For any $n \geq 0$ and any bounded function f, we have

$$\mathbb{E}(\gamma_n^N(f) 1_{\{N_n > 0\}}) = \gamma_n(f),$$

$$\sup_{f:\|f\|_\infty \leq 1} \mathbb{E}\left([1_{\{N_n > 0\}} \gamma_n^N(f) - \gamma_n(f)]^2 \right) \leq \frac{b^2(n)}{N_{inf}}.$$

Main Theorems: Law of Large Numbers

Theorem (Law of Large Numbers)

For any $n \geq 0$ and any bounded function f, we have

$$\mathbb{E}(\gamma_n^N(f)1_{\{N_n > 0\}}) = \gamma_n(f),$$

$$\sup_{f:\|f\|_\infty \leq 1} \mathbb{E}\left([1_{\{N_n > 0\}}\gamma_n^N(f) - \gamma_n(f)]^2\right) \leq \frac{b^2(n)}{N_{inf}}.$$

LLN for normalized measures

For any $n \geq 0$ and any bounded function f, we have

$$\sup_{f:\|f\|_\infty \leq 1} \left| \mathbb{E}\left[\eta_n^N(f)1_{\{N_n > 0\}} - \eta_n(f)\right] \right| \leq \frac{b(n)^2}{N_{\text{inf}}} + a(n)e^{-N_{\text{inf}}/c(n)}.$$

Main Theorems: Law of Large Numbers

Theorem (Law of Large Numbers)

For any $n \geq 0$ and any bounded function f, we have

$$\mathbb{E}(\gamma_n^N(f) 1_{\{N_n > 0\}}) = \gamma_n(f),$$

$$\sup_{f : \|f\|_\infty \leq 1} \mathbb{E}\left([1_{\{N_n > 0\}} \gamma_n^N(f) - \gamma_n(f)]^2 \right) \leq \frac{b^2(n)}{N_{inf}}.$$

LLN for normalized measures

For any $n \geq 0$ and any bounded function f, we have

$$\sup_{f : \|f\|_\infty \leq 1} \left| \mathbb{E}\left[\eta_n^N(f) 1_{\{N_n > 0\}} - \eta_n(f) \right] \right| \leq \frac{b(n)^2}{N_{\text{inf}}} + a(n) e^{-N_{\text{inf}}/c(n)}.$$

$$\sup_{f \,:\, \|f\|_\infty \leq 1} \left[\mathbb{E} \left| 1_{\{N_n > 0\}} \eta_n^N(f) - \eta_n(f) \right|^2 \right]^{1/2} \leq \frac{2b^2(n)}{|\gamma_n(1)|^2 N_{\text{inf}}}.$$

Plan

Main Theorems: Central Limit Theorem

Theorem (Central Limit Theorem)

Let $N \to \infty$ in such a way that $\rho_j = N^j/N$ are "preserved" for all $j \in \mathbb{M}$. Then, the random variable

$$\sqrt{N}\left(1_{\{N_{n+1} > 0\}}\gamma_{n+1}^N(1) - \mathbb{P}(T_n < T)\right)$$

converges in law to a Gaussian random variable with mean 0 and variance W_{n+1}, where

Main Theorems: Central Limit Theorem

Theorem (Central Limit Theorem)

Let $N \to \infty$ in such a way that $\rho_j = N^j/N$ are "preserved" for all $j \in \mathbb{M}$. Then, the random variable

$$\sqrt{N}\left(1_{\{N_{n+1} > 0\}}\gamma_{n+1}^N(1) - \mathbb{P}(T_n < T)\right)$$

converges in law to a Gaussian random variable with mean 0 and variance W_{n+1}, where

$$\frac{W_{n+1}}{\mathbb{P}^2(T_n < T)} = \sum_{q=0}^{n+1} \Omega_q \left(\frac{1}{P_q} - 1\right) + \sum_{q=0}^{n+1} \frac{\Omega_q}{P_q}\left[\frac{\widehat{\eta_q}\left([\Delta_q^n \circ \pi]^2\right)}{\widehat{\eta_q}^2\left(\Delta_q^n \circ \pi\right)} - 1\right],$$

with

Main Theorems: Central Limit Theorem

Theorem (Central Limit Theorem)

Let $N \to \infty$ in such a way that $\rho_j = N^j/N$ are "preserved" for all $j \in \mathbb{M}$. Then, the random variable

$$\sqrt{N}\left(1_{\{N_{n+1} > 0\}}\gamma_{n+1}^N(1) - \mathbb{P}(T_n < T)\right)$$

converges in law to a Gaussian random variable with mean 0 and variance W_{n+1}, where

$$\frac{W_{n+1}}{\mathbb{P}^2(T_n < T)} = \sum_{q=0}^{n+1}\Omega_q\left(\frac{1}{P_q} - 1\right) + \sum_{q=0}^{n+1}\frac{\Omega_q}{P_q}\left[\frac{\widehat{\eta}_q\left([\Delta_q^n \circ \pi]^2\right)}{\widehat{\eta}_q^{\,2}\left(\Delta_q^n \circ \pi\right)} - 1\right],$$

with

$$\Omega_q = \sum_{j \in \mathbb{M}}(\omega_{q-1}^j)^2\rho_j^{-1} = 1 + \chi^2(\omega_{q-1}, \rho),$$

and

$$\Delta_q^n(t, z) = \mathbb{P}(T_n \leq T | T_q = t, Z_{T_q} = z).$$

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:

CONCLUSION

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:
  - use importance sampling technique to make rare switches more frequent,

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:
  - use importance sampling technique to make rare switches more frequent,
  - or aggregate the modes in order to decrease the complexity (for large scale distributed hybrid systems).

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:
  - use importance sampling technique to make rare switches more frequent,
  - or aggregate the modes in order to decrease the complexity (for large scale distributed hybrid systems).
- Thus, we need to extend the previous results.

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:
  - use importance sampling technique to make rare switches more frequent,
  - or aggregate the modes in order to decrease the complexity (for large scale distributed hybrid systems).
- Thus, we need to extend the previous results.
- A better comprehension of the expression of W_{n+1} could help the choice of the N^j regarding the cost of the algorithm.

Conclusion

- The ”sampling per mode” algorithm has ”good” properties.
- However, to gain more time of simulation, we can:
  - use importance sampling technique to make rare switches more frequent,
  - or aggregate the modes in order to decrease the complexity (for large scale distributed hybrid systems).
- Thus, we need to extend the previous results.
- A better comprehension of the expression of W_{n+1} could help the choice of the N^j regarding the cost of the algorithm.
- This algorithm is implemented in a software developed by National Aerospace Laboratory (NLR) and used to evaluate the safety characteristics of an arbitrary (new) operational Air Traffic Management concept [Blom, 2009].

Références

Henk Blom, Bert Bakker, and Jaroslav Krystul.,
Rare event estimation for a large-scale stochastic hybrid system with air traffic application. *Rare Event Simulation using Monte Carlo Methods, pp. 194–214. Wiley, 2009*.

Krystul, J.,
Modelling of Stochastic Hybrid Systems with Applications to Accident Risk Assessment (2006). *PhD Dissertation: University of Twente*.

Del Moral, P. (2004).
Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications (2004). *Springer-Verlag*.

Le Gland, F. and Oudjane, N.,
A sequential particle algorithm that keeps the particle system alive. (2006) *in Stochastic Hybrid Systems : Theory and Safety Critical Applications, pp. 351-389, Springer, 2006*

F. Cérou, P. Del Moral, F. Le Gland and P. Lezaud,
Genetic genealogical models in rare event analysis, *ALEA, volume 1, paper 8, 2006*