

RARE EVENT SIMULATION USING HAMILTONIAN MONTE-CARLO AND SEQUENTIAL MONTE CARLO (SMC)

RISK FORUM 2018

Raphael Douady^{(1),(2),(3)}, **Shohruh Miryusupov**^{(1),(3),(4)}

⁽¹⁾ Labex ReFi, ⁽²⁾ CNRS,

⁽³⁾ Paris I, ⁽⁴⁾ MPG Partners)

Paris, 26-27 March 2018



Rare event simulation using Hamiltonian Monte-Carlo and Sequential Monte Carlo (SMC)

Introduction and state-of-the art of particle methods

Hamiltonian flow Monte Carlo methods

Application to barrier option pricing

Rare Event Probability Estimation

Given some stochastic process $\{X_t\}_{t \geq 0}$, for any test function f , compute a rare event probability: $P(f(X_T) \geq b) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{f(X_T) \geq b}]$.

- **Crude Monte Carlo (CMC):**

Simulate M realizations of a random variable $X^{(m)}$ and get the estimate:

$$P(f(X_T) \geq b) \approx \frac{1}{M} \sum_{m=1}^M \mathbb{1}_{f(X_T^{(m)}) \geq b} \quad (1)$$

⚠ A high number of samples is needed, especially with fat tailed distributions.

- **Refined approaches**

- Importance sampling (IS), large deviations theory, EVT, line sampling
- **Sequential Monte Carlo Methods (SMC)**
P Del Moral et al: Genealogical particle analysis of rare events
- **Sequential Markov Chain Monte Carlo Methods (SMCMC)**

Rare Event Probability Estimation

Given some stochastic process $\{X_t\}_{t \geq 0}$, for any test function f , compute a rare event probability: $P(f(X_T) \geq b) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{f(X_T) \geq b}]$.

- **Crude Monte Carlo (CMC):**

Simulate M realizations of a random variable $X^{(m)}$ and get the estimate:

$$P(f(X_T) \geq b) \approx \frac{1}{M} \sum_{m=1}^M \mathbb{1}_{f(X_T^{(m)}) \geq b} \quad (1)$$

⚠ A high number of samples is needed, especially with fat tailed distributions.

- **Refined approaches**

- Importance sampling (IS), large deviations theory, EVT, line sampling
- **Sequential Monte Carlo Methods (SMC)**
P. Del Moral et al.: Genealogical particle analysis of rare events
- **Sequential Markov Chain Monte Carlo Methods (SMCMC)**

Rare Event Probability Estimation

Given some stochastic process $\{X_t\}_{t \geq 0}$, for any test function f , compute a rare event probability: $P(f(X_T) \geq b) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{f(X_T) \geq b}]$.

- **Crude Monte Carlo (CMC):**

Simulate M realizations of a random variable $X^{(m)}$ and get the estimate:

$$P(f(X_T) \geq b) \approx \frac{1}{M} \sum_{m=1}^M \mathbb{1}_{f(X_T^{(m)}) \geq b} \quad (1)$$

⚠ A high number of samples is needed, especially with fat tailed distributions.

- **Refined approaches**

- Importance sampling (IS), large deviations theory, EVT, line sampling
- **Sequential Monte Carlo Methods (SMC)**
P. Del Moral et al.: Genealogical particle analysis of rare events
- **Sequential Markov Chain Monte Carlo Methods (SMCMC)**

Rare Event Probability Estimation

Given some stochastic process $\{X_t\}_{t \geq 0}$, for any test function f , compute a rare event probability: $P(f(X_T) \geq b) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{f(X_T) \geq b}]$.

- **Crude Monte Carlo (CMC):**

Simulate M realizations of a random variable $X^{(m)}$ and get the estimate:

$$P(f(X_T) \geq b) \approx \frac{1}{M} \sum_{m=1}^M \mathbb{1}_{f(X_T^{(m)}) \geq b} \quad (1)$$

⚠ A high number of samples is needed, especially with fat tailed distributions.

- **Refined approaches**

- Importance sampling (IS), large deviations theory, EVT, line sampling
- **Sequential Monte Carlo Methods (SMC)**
P. Del Moral et al.: Genealogical particle analysis of rare events
- **Sequential Markov Chain Monte Carlo Methods (SMCMC)**

Rare Event Probability Estimation

Given some stochastic process $\{X_t\}_{t \geq 0}$, for any test function f , compute a rare event probability: $P(f(X_T) \geq b) = \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{f(X_T) \geq b}]$.

- **Crude Monte Carlo (CMC):**

Simulate M realizations of a random variable $X^{(m)}$ and get the estimate:

$$P(f(X_T) \geq b) \approx \frac{1}{M} \sum_{m=1}^M \mathbb{1}_{f(X_T^{(m)}) \geq b} \quad (1)$$

⚠ A high number of samples is needed, especially with fat tailed distributions.

- **Refined approaches**

- Importance sampling (IS), large deviations theory, EVT, line sampling
- **Sequential Monte Carlo Methods (SMC)**
P. Del Moral et al.: Genealogical particle analysis of rare events
- **Sequential Markov Chain Monte Carlo Methods (SMCMC)**

Feynman-Kac Measures

A sequence of Feynman - Kac measures is defined for any test function f as:

$$\eta_n f = \frac{\int \dots \int f(x_n) \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})}{\int \dots \int \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})} \quad (n \in \mathbb{N})$$

- where $\{\mathcal{M}_n\}_{n \in \mathbb{N}}$ the set of unnormalized Markov transitions on some measurable state space (X, \mathcal{X}) ;
- η_0 is initial probability measure;
- $\eta_n^M f \xrightarrow{M \rightarrow \infty} \eta_n f$ (a.s) and $\sqrt{N}(\eta_n^M f - \eta_n f) \xrightarrow[M \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(f))$
 where $\eta_n^M f = \sum_{m=1}^M \frac{\omega_n^{(m)} f(\zeta_n^{(m)})}{\sum_{j=1}^M \omega_n^{(j)}}$.

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

Douc, R. and Moulines, E., Limit theorems for weighted samples with applications to SMC methods.(2008)

Feynman-Kac Measures

A sequence of Feynman - Kac measures is defined for any test function f as:

$$\eta_n f = \frac{\int \dots \int f(x_n) \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})}{\int \dots \int \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})} \quad (n \in \mathbb{N})$$

- where $\{\mathcal{M}_n\}_{n \in \mathbb{N}}$ the set of unnormalized Markov transitions on some measurable state space (X, \mathcal{X}) ;
- η_0 is initial probability measure;
- $\eta_n^M f \xrightarrow{M \rightarrow \infty} \eta_n f$ (a.s) and $\sqrt{N}(\eta_n^M f - \eta_n f) \xrightarrow[M \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(f))$
 where $\eta_n^M f = \sum_{m=1}^M \frac{\omega_n^{(m)} f(\xi_n^{(m)})}{\sum_{j=1}^M \omega_n^{(j)}}$.

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

Douc, R. and Moulines, E., Limit theorems for weighted samples with applications to SMC methods.(2008)

Feynman-Kac Measures

A sequence of Feynman - Kac measures is defined for any test function f as:

$$\eta_n f = \frac{\int \dots \int f(x_n) \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})}{\int \dots \int \eta_0(dx_0) \prod_{i=1}^{n-1} \mathcal{M}_i(x_i, dx_{i+1})} \quad (n \in \mathbb{N})$$

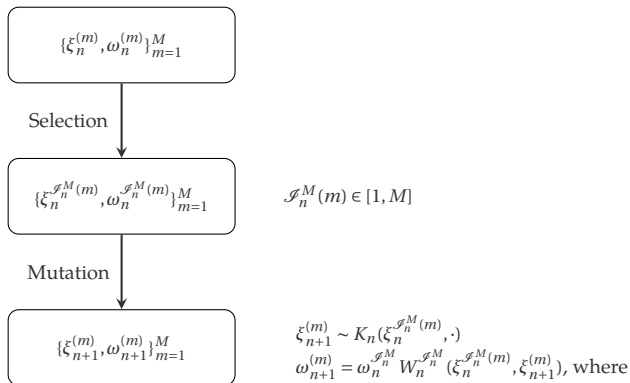
- where $\{\mathcal{M}_n\}_{n \in \mathbb{N}}$ the set of unnormalized Markov transitions on some measurable state space (X, \mathcal{X}) ;
- η_0 is initial probability measure;
- $\eta_n^M f \xrightarrow{M \rightarrow \infty} \eta_n f$ (a.s) and $\sqrt{N}(\eta_n^M f - \eta_n f) \xrightarrow[M \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma(f))$
 where $\eta_n^M f = \sum_{m=1}^M \frac{\omega_n^{(m)} f(\xi_n^{(m)})}{\sum_{j=1}^M \omega_n^{(j)}}$.

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

Douc, R. and Moulines, E., Limit theorems for weighted samples with applications to SMC methods.(2008)

Sequential Monte Carlo

SMC allows to approximate the empirical measure by a set of weighted particles ξ, ω .



- $K_n(\xi_n, \cdot)$ is the transition kernel
- W_n is the importance weight function such that

$$W_n(x, y) = \frac{Q_n(x, dy)}{K_n(x, dy)} \quad (x, y) \in \mathcal{X}^2$$

Sequential Importance Resampling/Bootstrap Resampling

```

1 Initialization:  $M$  - #(simulations),  $N$  - #(time steps),  $\eta_0$  - initial measure
2 for  $m = 1, \dots, M$  do
3    $\xi_0^{(m)} \sim \eta_0$ ;
4    $\xi_1^{(m)} \sim K_0(\xi_0^{(m)}, \cdot)$ ;
5    $\omega_1^{(m)} = \omega_0^{(m)}(\xi_0^{(m)}, \xi_1^{(m)})$ 
6 end
7 for  $n = 1, \dots, N$  do
8   for  $m = 1, \dots, M$  do
9     if  $n < N$  then
10      Selection step:
11      Sample indices  $\mathcal{J}_{n-1}^M(m)$  according to sampling strategy  $\mathcal{S}(\{\omega_{n-1}^{(m)}\}_{m=1}^M)$ ;
12      end
13      Mutation step:
14      Generate  $\xi_n^{(m)}$  from  $K_n(\cdot, \xi_{n-1}^{\mathcal{J}_{n-1}^M(m)})$  and set  $\widehat{\xi}_n^{(m)} = (\widehat{\xi}_n^{(m)}, \xi_{n-1}^{(m)})$ ;
15      Compute the weight:  $\omega_n^{(m)}(\widehat{\xi}_n^{(m)}) = W_{n-1}(\xi_{n-1}^{\mathcal{J}_{n-1}^M(m)}, \xi_n^{\mathcal{J}_{n-1}^M(m)})$ ;
16    end
17 end

```

Markov Chain Monte Carlo – Metropolis-Hastings Algorithm

MCMC: Given a target π , design a transition kernel \mathcal{K} such that asymptotically:

$$\frac{1}{N} \sum_{j=1}^N f(X^{(j)}) \xrightarrow{n \rightarrow \infty} \int_{\mathcal{X}} f(x) \pi(x) dx \quad \text{and} \quad X^{(j)} \sim \pi \quad (3)$$

Introduce a family of proposal distribution $\{q(x, \cdot), x \in \mathcal{X}\}$, i.e. $\int q(x, y) dy = 1$ for any $x \in \mathcal{X}$.

The main idea of MH MCMC is,

- given that the Markov Chain is $x \in \mathcal{X}$, to propose a new candidate from $q(x, \cdot)$,
- to accept new proposal with an acceptance probability $a(x, y)$, which insures that the invariant distribution of the transition kernel \mathcal{K} is the target distribution $\pi(x)$

Markov Chain Monte Carlo – Metropolis-Hastings Algorithm

MCMC: Given a target π , design a transition kernel \mathcal{K} such that asymptotically:

$$\frac{1}{N} \sum_{j=1}^N f(X^{(j)}) \xrightarrow{n \rightarrow \infty} \int_{\mathcal{X}} f(x) \pi(x) dx \quad \text{and} \quad X^{(j)} \sim \pi \quad (3)$$

Introduce a family of proposal distribution $\{q(x, \cdot), x \in \mathcal{X}\}$, i.e. $\int q(x, y) dy = 1$ for any $x \in \mathcal{X}$.

The main idea of MH MCMC is,

- given that the Markov Chain is $x \in \mathcal{X}$, to propose a new candidate from $q(x, \cdot)$,
- to accept new proposal with an acceptance probability $a(x, y)$, which insures that the invariant distribution of the transition kernel \mathcal{K} is the target distribution $\pi(x)$

"Sequential" Markov Chain Monte Carlo/MH

-
-
- 1 Initialization: N - #(time steps);
 - 2 $X_{0,0}^j \sim \mathcal{K}_0(X_{0,0}^{j-1}, \cdot)$ with \mathcal{K}_0 is a π_0 invariant MCMC kernel;
 - 3 **for** $n = 1, \dots, N$ **do**
 - 4 **for** $j = 1, \dots, M + M_b$ **do**
 - 5 | Sample $X_{n,0:n}^j = \mathcal{K}_n(X_{n,0:n}^{j-1}, \cdot)$ with \mathcal{K}_n is a π_n invariant MCMC kernel;
 - 6 **end**
 - 7 $\pi(x_{0:n}) \approx \frac{1}{M} \sum_{j=M_b+1}^{M_b+M} \delta_{X_{n,0:n}^j}(dx_{0:n})$
 - 8 **end**
-

Algorithm1: MH algorithm for the kernel \mathcal{K}

- 1 Input: $X_{n,0:n}^{j-1}$
 - 2 Compute the acceptance probability: $a(X_{n,0:n}, x^*) = 1 \wedge \frac{\pi(x^*)q(x^*, X_{n,0:n}^{j-1})}{\pi(X_{n,0:n}^{j-1})q(X_{n,0:n}, x^*)}$
 - 3 assuming that $a(x, y) = 0$ if $\pi(x)q(x, y) = 0$;
 - 4 Draw $u \sim \mathcal{U}(0, 1)$ and set $X_{n,0:n}^j = x^*$, if $u < a$, and $X_{n,0:n}^j = X_{n,0:n}^{j-1}$ otherwise;
-

Sequential MCMC

Problem: Slow convergence to the invariant density, when naive MCMC MH kernels are used.

Solution: Use Metropolized Hamiltonian system's proposals as MCMC kernel by augmenting r.v. with an auxiliary variables

- Recipe comes from statistical physics and mathematical mechanics
- Hamiltonian flow's symplecticness and energy-conservation property entails ergodicity
- We replace the exact flow by its numerical discretization, however, invariant measure is no longer preserved under the discrete flow.
- We "Metropolize" to enforce the desired target distribution thereby accounting for the energy error of the numerical flow and to have an invariant measure.
- Geometrically ergodic for regular potentials

Duane, S, Kennedy, AD, Pendleton, BJ, and Roweth, D. Hybrid Monte Carlo. Physics letters B, 1987.

Sequential MCMC

Problem: Slow convergence to the invariant density, when naive MCMC MH kernels are used.

Solution: Use Metropolized Hamiltonian system's proposals as MCMC kernel by augmenting r.v. with an auxiliary variables

- Recipe comes from statistical physics and mathematical mechanics
- Hamiltonian flow's symplecticness and energy-conservation property entails ergodicity
- We replace the exact flow by its numerical discretization, however, invariant measure is no longer preserved under the discrete flow.
- We "Metropolize" to enforce the desired target distribution thereby accounting for the energy error of the numerical flow and to have an invariant measure.
- Geometrically ergodic for regular potentials

Duane, S, Kennedy, AD, Pendleton, BJ, and Roweth, D. Hybrid Monte Carlo. Physics letters B, 1987.

Sequential MCMC

Problem: Slow convergence to the invariant density, when naive MCMC MH kernels are used.

Solution: Use Metropolized Hamiltonian system's proposals as MCMC kernel by augmenting r.v. with an auxiliary variables

- Recipe comes from statistical physics and mathematical mechanics
- Hamiltonian flow's symplecticness and energy-conservation property entails ergodicity
- We replace the exact flow by its numerical discretization, however, invariant measure is no longer preserved under the discrete flow.
- We "Metropolize" to enforce the desired target distribution thereby accounting for the energy error of the numerical flow and to have an invariant measure.
- Geometrically ergodic for regular potentials

Duane, S, Kennedy, AD, Pendleton, BJ, and Roweth, D. Hybrid Monte Carlo. Physics letters B, 1987.

Sequential MCMC

Problem: Slow convergence to the invariant density, when naive MCMC MH kernels are used.

Solution: Use Metropolized Hamiltonian system's proposals as MCMC kernel by augmenting r.v. with an auxiliary variables

- Recipe comes from statistical physics and mathematical mechanics
- Hamiltonian flow's symplecticness and energy-conservation property entails ergodicity
- We replace the exact flow by its numerical discretization, however, invariant measure is no longer preserved under the discrete flow.
- We "Metropolize" to enforce the desired target distribution thereby accounting for the energy error of the numerical flow and to have an invariant measure.
- Geometrically ergodic for regular potentials

Duane, S, Kennedy, AD, Pendleton, BJ, and Roweth, D. Hybrid Monte Carlo. Physics letters B, 1987.

Sequential MCMC

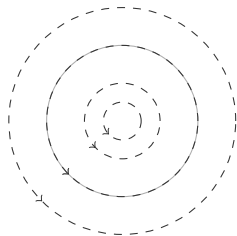
Problem: Slow convergence to the invariant density, when naive MCMC MH kernels are used.

Solution: Use Metropolized Hamiltonian system's proposals as MCMC kernel by augmenting r.v. with an auxiliary variables

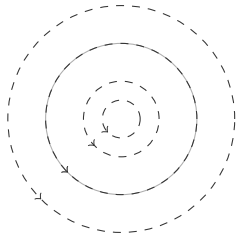
- Recipe comes from statistical physics and mathematical mechanics
- Hamiltonian flow's symplecticness and energy-conservation property entails ergodicity
- We replace the exact flow by its numerical discretization, however, invariant measure is no longer preserved under the discrete flow.
- We "Metropolize" to enforce the desired target distribution thereby accounting for the energy error of the numerical flow and to have an invariant measure.
- Geometrically ergodic for regular potentials

Duane, S, Kennedy, AD, Pendleton, BJ, and Roweth, D. Hybrid Monte Carlo. *Physics letters B*, 1987.

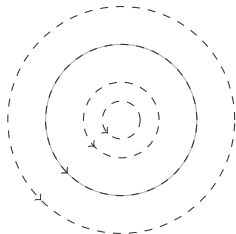
HMC Intro



HMC Intro

 $X \rightarrow (X, P)$

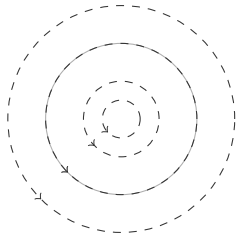
HMC Intro



$$X \rightarrow (X, P)$$

$$\mathcal{H}(X, P) \rightarrow e^{-\mathcal{H}(X, P)} d^n P d^n X$$

HMC Intro

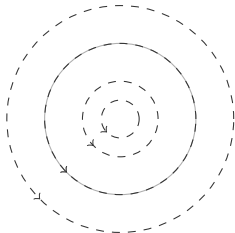


$$X \rightarrow (X, P)$$

$$\mathcal{H}(X, P) \rightarrow e^{-\mathcal{H}(X, P)} d^n P d^n X$$

$$\mathcal{H}(X, P) = -\log \pi(X, P)$$

HMC Intro



$$X \rightarrow (X, P)$$

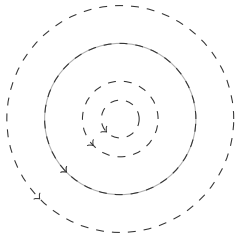
$$\mathcal{H}(X, P) \rightarrow e^{-\mathcal{H}(X, P)} d^n P d^n X$$

$$\mathcal{H}(X, P) = -\log \pi(X, P)$$



$$\mathcal{H}(X, P) = -\log \pi(P|X)\pi(X)$$

HMC Intro



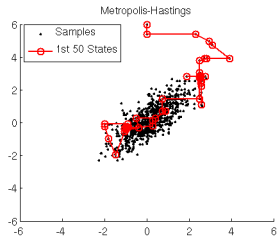
$$X \rightarrow (X, P)$$

$$\mathcal{H}(X, P) \rightarrow e^{-\mathcal{H}(X, P)} d^n P d^n X$$

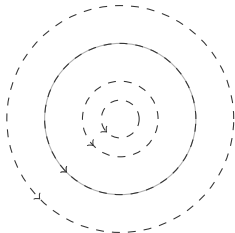
$$\mathcal{H}(X, P) = -\log \pi(X, P)$$



$$\mathcal{H}(X, P) = -\log \pi(P|X)\pi(X)$$



HMC Intro



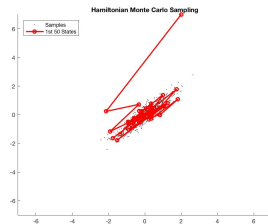
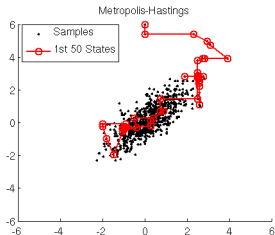
$$X \rightarrow (X, P)$$

$$\mathcal{H}(X, P) \rightarrow e^{-\mathcal{H}(X, P)} d^n P d^n X$$

$$\mathcal{H}(X, P) = -\log \pi(X, P)$$



$$\mathcal{H}(X, P) = -\log \pi(P|X)\pi(X)$$



HMC. Example

- MCMC is useful when sampling a r.v. X having target density known up to a normalizing const:

$$\pi(X) = \frac{e^{-\Psi(X)}}{\mathcal{Z}} \quad \text{where} \quad (4)$$

- Consider a Gaussian log-likelihood function:

$$\log p(x_n | \theta) \propto \left(-\frac{1}{2} \left(\frac{x_n - k(x_{n-1}, \theta)}{\sigma} \right)^2 \right) = -\Psi(X) \quad (5)$$

- Extend the state space by a momentum r.v. P , so that the joint-distribution is given by:

$$\pi(X, P) = e^{-\mathcal{H}(X, P)} \propto e^{-\Psi(X)} e^{-\frac{1}{2} P^T \mathfrak{M} P} \quad (6)$$

- Transition step:** Solve the following ODE with initial conditions:

$$\begin{aligned} \partial_u P_u &= -\nabla_x \Psi(X) & P_0 &\sim e^{\frac{1}{2} P^T \mathfrak{M} P} & P^* &= P_L \\ \partial_u X_u &= P & X_0 &= X_0 & x^* &= X_L \end{aligned} \quad (7)$$

- Acceptance step:** accept x^*, P^* with probability:

$$a = \min \left(1, \frac{\pi(x^*, P^*)}{\pi(X, P)} \right) = 1 \wedge e^{-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X, P))} \quad (8)$$

HMC. Example

- MCMC is useful when sampling a r.v. X having target density known up to a normalizing const:

$$\pi(X) = \frac{e^{-\Psi(X)}}{\mathcal{Z}} \quad \text{where} \quad (4)$$

- Consider a Gaussian log-likelihood function:

$$\log p(x_n | \theta) \propto \left(-\frac{1}{2} \left(\frac{x_n - k(x_{n-1}, \theta)}{\sigma} \right)^2 \right) = -\Psi(X) \quad (5)$$

- Extend the state space by a momentum r.v. P , so that the joint-distribution is given by:

$$\pi(X, P) = e^{-\mathcal{H}(X, P)} \propto e^{-\Psi(X)} e^{-\frac{1}{2} P^T \mathfrak{M} P} \quad (6)$$

- Transition step:** Solve the following ODE with initial conditions:

$$\begin{aligned} \partial_u P_u &= -\nabla_x \Psi(X) & P_0 &\sim e^{\frac{1}{2} P^T \mathfrak{M} P} & P^* &= P_L \\ \partial_u X_u &= P & X_0 &= X_0 & x^* &= X_L \end{aligned} \quad (7)$$

- Acceptance step:** accept x^*, P^* with probability:

$$a = \min \left(1, \frac{\pi(x^*, P^*)}{\pi(X, P)} \right) = 1 \wedge e^{-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X, P))} \quad (8)$$

HMC. Example

- MCMC is useful when sampling a r.v. X having target density known up to a normalizing const:

$$\pi(X) = \frac{e^{-\Psi(X)}}{\mathcal{Z}} \quad \text{where} \quad (4)$$

- Consider a Gaussian log-likelihood function:

$$\log p(x_n|\theta) \propto \left(-\frac{1}{2} \left(\frac{x_n - k(x_{n-1}, \theta)}{\sigma} \right)^2 \right) = -\Psi(X) \quad (5)$$

- Extend the state space by a momentum r.v. P , so that the joint-distribution is given by:

$$\pi(X, P) = e^{-\mathcal{H}(X, P)} \propto e^{-\Psi(X)} e^{-\frac{1}{2} P^T \mathfrak{M} P} \quad (6)$$

- Transition step:** Solve the following ODE with initial conditions:

$$\begin{aligned} \partial_u P_u &= -\nabla_x \Psi(X) & P_0 &\sim e^{\frac{1}{2} P^T \mathfrak{M} P} & P_* &= P_L \\ \partial_u X_u &= P & X_0 &= X_0 & x^* &= X_L \end{aligned} \quad (7)$$

- Acceptance step:** accept x^*, P^* with probability:

$$a = \min \left(1, \frac{\pi(x^*, P^*)}{\pi(X, P)} \right) = 1 \wedge e^{-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X, P))} \quad (8)$$

HMC. Example

- MCMC is useful when sampling a r.v. X having target density known up to a normalizing const:

$$\pi(X) = \frac{e^{-\Psi(X)}}{\mathcal{Z}} \quad \text{where} \quad (4)$$

- Consider a Gaussian log-likelihood function:

$$\log p(x_n|\theta) \propto \left(-\frac{1}{2} \left(\frac{x_n - k(x_{n-1}, \theta)}{\sigma} \right)^2 \right) = -\Psi(X) \quad (5)$$

- Extend the state space by a momentum r.v. P , so that the joint-distribution is given by:

$$\pi(X, P) = e^{-\mathcal{H}(X, P)} \propto e^{-\Psi(X)} e^{-\frac{1}{2} P^T \mathfrak{M} P} \quad (6)$$

- Transition step:** Solve the following ODE with initial conditions:

$$\begin{aligned} \partial_u P_u &= -\nabla_x \Psi(X) & P_0 &\sim e^{\frac{1}{2} P^T \mathfrak{M} P} & P^* &= P_L \\ \partial_u X_u &= P & X_0 &= X_0 & x^* &= X_L \end{aligned} \quad (7)$$

- Acceptance step:** accept x^*, P^* with probability:

$$a = \min \left(1, \frac{\pi(x^*, P^*)}{\pi(X, P)} \right) = 1 \wedge e^{-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X, P))} \quad (8)$$

HMC. Example

- MCMC is useful when sampling a r.v. X having target density known up to a normalizing const:

$$\pi(X) = \frac{e^{-\Psi(X)}}{\mathcal{Z}} \quad \text{where} \quad (4)$$

- Consider a Gaussian log-likelihood function:

$$\log p(x_n|\theta) \propto \left(-\frac{1}{2} \left(\frac{x_n - k(x_{n-1}, \theta)}{\sigma} \right)^2 \right) = -\Psi(X) \quad (5)$$

- Extend the state space by a momentum r.v. P , so that the joint-distribution is given by:

$$\pi(X, P) = e^{-\mathcal{H}(X, P)} \propto e^{-\Psi(X)} e^{-\frac{1}{2} P^T \mathfrak{M} P} \quad (6)$$

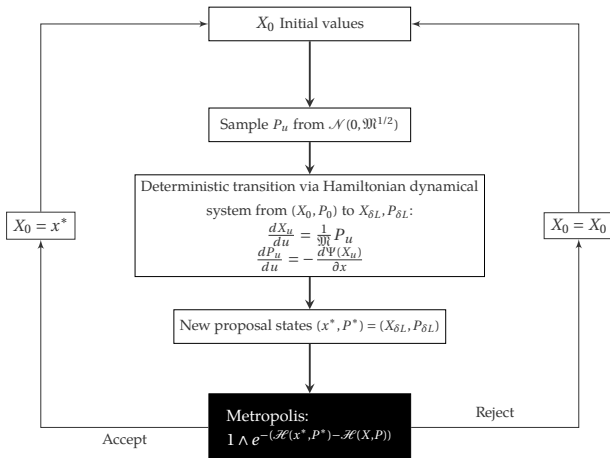
- Transition step:** Solve the following ODE with initial conditions:

$$\begin{aligned} \partial_u P_u &= -\nabla_x \Psi(X) & P_0 &\sim e^{\frac{1}{2} P^T \mathfrak{M} P} & P^* &= P_L \\ \partial_u X_u &= P & X_0 &= X_0 & x^* &= X_L \end{aligned} \quad (7)$$

- Acceptance step:** accept x^*, P^* with probability:

$$a = \min \left(1, \frac{\pi(x^*, P^*)}{\pi(X, P)} \right) = 1 \wedge e^{-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X, P))} \quad (8)$$

Hamiltonian Flow Monte Carlo Scheme



HMC in Rare Event Setting

- We compute the loglikelihood and its derivative, sample momentum variables
- Generate new proposal x^* by simulating Hamiltonian dynamics;
- Accept new proposals $X_{n+1} = x^*, P_{n+1} = P^*$ if acceptance probability satisfies $a > \mathcal{U}ni f(0, 1)$ and $x^* \in A$:

$$a = \min(1, \exp(-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X_n, P_n))))$$

HMC in Rare Event Setting

- We compute the loglikelihood and its derivative, sample momentum variables
- Generate new proposal x^* by simulating Hamiltonian dynamics;
- Accept new proposals $X_{n+1} = x^*, P_{n+1} = P^*$ if acceptance probability satisfies $a > \mathcal{U}ni f(0, 1)$ and $x^* \in A$:

$$a = \min(1, \exp(-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X_n, P_n))))$$

HMC in Rare Event Setting

- We compute the loglikelihood and its derivative, sample momentum variables
- Generate new proposal x^* by simulating Hamiltonian dynamics;
- Accept new proposals $X_{n+1} = x^*, P_{n+1} = P^*$ if acceptance probability satisfies $a > \mathcal{U}ni f(0, 1)$ and $x^* \in A$:
 - $a = \min(1, \exp(-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X_n, P_n))))$,

HMC in Rare Event Setting

- We compute the loglikelihood and its derivative, sample momentum variables
- Generate new proposal x^* by simulating Hamiltonian dynamics;
- Accept new proposals $X_{n+1} = x^*, P_{n+1} = P^*$ if acceptance probability satisfies $a > \mathcal{U}ni f(0, 1)$ and $x^* \in A$:
 - $a = \min(1, \exp(-(\mathcal{H}(x^*, P^*) - \mathcal{H}(X_n, P_n))))$,

Hamiltonian Monte Carlo Algorithm

```

1 Initialization:  $n_S$  - #(simulations),  $n_t$  - #(time steps)
2 for  $n = 1, \dots, n_t$  do
3   for  $m = 1, \dots, n_S$  do
4     Generate  $X_n^{(m)}$  from prior  $\tilde{p}(X_0^{(m)}, \cdot)$ ;
5     Simulate initial momentum  $P_1^{(m)} \sim \mathcal{N}(0, I_M)$ , set  $x_H^{(m)} = X_n^{(m)}$  and run Hamiltonian
     flow:
6     for  $l_f = 1, \dots, L-1$  do
7        $P_H^{(m)}((l_f + \frac{1}{2})\delta) = P_H^{(m)}(l_f) - \frac{\delta}{2} \frac{\partial \Psi}{\partial x_H}(x_H^{(m)}(l_f))$ 
7        $x_H^{(m)}((l_f + 1)\delta) = x_H^{(m)}(l_f) + \delta P_H^{(m)}((l_f + \frac{1}{2})\delta) I_M^{-1}$ 
7        $P_H^{(m)}((l_f + 1)\delta) = P_H^{(m)}((l_f + \frac{1}{2})\delta) - \frac{\delta}{2} \frac{\partial \Psi}{\partial x_H}(x_H^{(m)}((l_f + 1)\delta))$ 
8     end
9     Set  $x^* = x_H^{(m)}(L)$ ,  $P^* = P_H^{(m)}(L)$  and calculate acceptance probability:
9      $a = 1 \wedge e^{(-\mathcal{H}(x^*, P^*) + \mathcal{H}(x_H^{(m)}, P_H^{(m)}))\Delta t}$ ;
10    Draw  $u \sim \mathcal{U}nif(0, 1)$  and set  $X_n^j = x^*$ , if  $u < a$  and  $x^* \in A$ , and  $X_n^j = X_n^{j-1}$  otherwise;
11  end
12 end
13 Compute:  $\hat{C}^{HFMC} = \frac{1}{n_S} \sum_{m=1}^{n_S} \left( f(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, P_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, P_n^{(m)})\Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right)$ 

```

Application of HMC on a barrier option pricing



○

$$C = e^{-(r-q)T} \mathbb{E}[g(X_{n_S}) \prod_{n=1}^{n_t} \mathbb{1}_{X_t \in [t_{n-1}, t_n] \in A_n}] \quad (9)$$

○

$$A_n = \inf_{t_{n-1} \leq t \leq t_n} \{t : X_t > B\}$$

○ HPMC estimator to compute DOC call option is given by:

$$\widehat{C}^{HPMC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, p_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, p_n^{(m)}) \Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right) \quad (10)$$

○ Monte Carlo estimate is given by:

$$\widehat{C}^{MC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} \mathbb{1}_{X_n^{(m)} \in A_n} \right) \quad (11)$$

○ The IPS estimator is given by:

$$\widehat{C}^{IPS} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(\widehat{X}_{n_t}^{(m)}) \prod_{t=1}^{n_t} W_{n-1}(\mathcal{X}_{n-1}^m) \mathbb{1}_{\widehat{X}_n \in A_n} \right) \quad (12)$$

Application of HMC on a barrier option pricing



○

$$C = e^{-(r-q)T} \mathbb{E}[g(X_{n_S}) \prod_{n=1}^{n_t} \mathbb{1}_{X_t \in [t_{n-1}, t_n] \in A_n}] \quad (9)$$

○

$$A_n = \inf_{t_{n-1} \leq t \leq t_n} \{t : X_t > B\}$$

○ HFMC estimator to compute DOC call option is given by:

$$\widehat{C}^{HFMC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, P_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, P_n^{(m)}) \Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right) \quad (10)$$

○ Monte Carlo estimate is given by:

$$\widehat{C}^{MC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} \mathbb{1}_{X_n^{(m)} \in A_n} \right) \quad (11)$$

○ The IPS estimator is given by:

$$\widehat{C}^{IPS} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(\widehat{X}_{n_t}^{(m)}) \prod_{t=1}^{n_t} W_{n-1}(\mathcal{X}_{n-1}^m) \mathbb{1}_{\widehat{X}_n \in A_n} \right) \quad (12)$$

Application of HMC on a barrier option pricing



○

$$C = e^{-(r-q)T} \mathbb{E}[g(X_{n_S}) \prod_{n=1}^{n_t} \mathbb{1}_{X_t \in [t_{n-1}, t_n] \in A_n}] \quad (9)$$

○

$$A_n = \inf_{t_{n-1} \leq t \leq t_n} \{t : X_t > B\}$$

○ HFMC estimator to compute DOC call option is given by:

$$\widehat{C}^{HFMC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, P_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, P_n^{(m)}) \Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right) \quad (10)$$

○ Monte Carlo estimate is given by:

$$\widehat{C}^{MC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} \mathbb{1}_{X_n^{(m)} \in A_n} \right) \quad (11)$$

○ The IPS estimator is given by:

$$\widehat{C}^{IPS} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(\widehat{X}_{n_t}^{(m)}) \prod_{i=1}^{n_t} W_{n-1}(\mathcal{X}_{n-1}^m) \mathbb{1}_{\widehat{X}_n \in A_n} \right) \quad (12)$$

Application of HMC on a barrier option pricing



○

$$C = e^{-(r-q)T} \mathbb{E}[g(X_{n_S}) \prod_{n=1}^{n_t} \mathbb{1}_{X_{t \in [t_{n-1}, t_n]} \in A_n}] \quad (9)$$

○

$$A_n = \inf_{t_{n-1} \leq t \leq t_n} \{t : X_t > B\}$$

○ HFMC estimator to compute DOC call option is given by:

$$\widehat{C}^{HFMC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, P_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, P_n^{(m)}) \Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right) \quad (10)$$

○ Monte Carlo estimate is given by:

$$\widehat{C}^{MC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} \mathbb{1}_{X_n^{(m)} \in A_n} \right) \quad (11)$$

○ The IPS estimator is given by:

$$\widehat{C}^{IPS} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(\widehat{X}_{n_t}^{(m)}) \prod_{l=1}^{n_t} W_{n-1}(\mathbf{X}_{n-1}^m) \mathbb{1}_{\widehat{X}_n \in A_n} \right) \quad (12)$$

Application of HMC on a barrier option pricing



○

$$C = e^{-(r-q)T} \mathbb{E} [g(X_{n_S}) \prod_{n=1}^{n_t} \mathbb{1}_{X_{t \in [t_{n-1}, t_n]} \in A_n}] \quad (9)$$

○

$$A_n = \inf_{t_{n-1} \leq t \leq t_n} \{t : X_t > B\}$$

○ HFMC estimator to compute DOC call option is given by:

$$\widehat{C}^{HFMC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} e^{(-\mathcal{H}(X_{n+1}^{(m)}, P_{n+1}^{(m)}) + \mathcal{H}(X_n^{(m)}, P_n^{(m)}) \Delta t} \mathbb{1}_{X_{n+1}^{(m)} \in A_n} \right) \quad (10)$$

○ Monte Carlo estimate is given by:

$$\widehat{C}^{MC} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(X_{n_t}^{(m)}) \prod_{n=1}^{n_t} \mathbb{1}_{X_n^{(m)} \in A_n} \right) \quad (11)$$

○ The IPS estimator is given by:

$$\widehat{C}^{IPS} = e^{-(r-q)T} \frac{1}{n_S} \sum_{m=1}^{n_S} \left(g(\widehat{X}_{n_t}^{(m)}) \prod_{l=1}^{n_t} W_{n-1}(\mathbf{X}_{n-1}^m) \mathbb{1}_{\widehat{X}_n \in A_n} \right) \quad (12)$$

Application of HMC on a barrier option pricing, $B = 65$

Table: DOC Barrier option estimates statistics. $X_0 = 100$, $K = 100$, $r = 0.1$, $\sigma = 0.3$, $T = 1/2$, and $div = 0$; $\delta = 0.0001$, $B = 65$, #(Leap frog step): 40. True price: 10.9064, $N = 75000$, $M = 750$

Stat	MC	IPS	HFMC
St. dev.	0.062385996	0.044259477	0.038039517
RMSE	0.037561882	0.051285344	0.037561882
RRMSE	0.000355199	0.000240548	0.000129293
CPU time	2.2626	6.0322	7.6832
FOM	13475.2	10117.7	10711.0

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795-808.

- Parallelization and convergence of HMC

- Concentration inequalities (The probability of having error)

- Information geometry (Sampling on statistical manifolds)

Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*

- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \quad (13)$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795-808.

- Parallelization and convergence of HMC
- Concentration inequalities (The probability of having error)
- Information geometry (Sampling on statistical manifolds)
Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*
- Adapt the algorithm for a sampling in general Hilbert spaces
Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808.

- Parallelization and convergence of HMC

- Concentration inequalities (The probability of having error)

- Information geometry (Sampling on statistical manifolds)

Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*

- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795-808.

- Parallelization and convergence of HMC

- Concentration inequalities (The probability of having error)

- Information geometry (Sampling on statistical manifolds)

Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*

- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808.

- Parallelization and convergence of HMC
- Concentration inequalities (The probability of having error)
- Information geometry (Sampling on statistical manifolds)

Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*

- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808.

- Parallelization and convergence of HMC
- Concentration inequalities (The probability of having error)
- Information geometry (Sampling on statistical manifolds)

Shun'ichi Amari (1985) *Differential-geometrical methods in statistics*

- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) *Advanced MCMC methods for sampling on diffusion pathspace.*

Further research and improvements

- Combine with splitting algorithm

Cerou, F. and Guyader, A. (2007). Adaptive multilevel splitting for rare event analysis.

- Estimate parameters of the rare event set using Particle MCMC:

$$\mathbb{P}(\Theta | f(x) > a) \tag{13}$$

Andrieu, C. et al., (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*.

- Use the Hamiltonian Kernel in SMC framework and study adapted particle methods.

Cerou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795-808.

- Parallelization and convergence of HMC
- Concentration inequalities (The probability of having error)
- Information geometry (Sampling on statistical manifolds)
- Adapt the algorithm for a sampling in general Hilbert spaces

Beskos, A; Kalogeropoulos, K; Pazos, E; (2013) Advanced MCMC methods for sampling on diffusion pathspace.

Merci a tous pour votre attention !
Thanks for your attention!