Long time stability of Feynman-Kac models.

Les probabilités de demain

Grégoire Ferré, Gabriel Stoltz, Mathias Rousset

CERMICS - ENPC, INRIA & Labex BEZOUT

gregoire.ferre@enpc.fr

Thursday, May 3rd, 2018

Outline

1. Introduction: physical motivations

2. Markov chains and ergodicity

3. Feynman–Kac dynamics

1. Introduction: physical motivations

Markov chains and physical systems

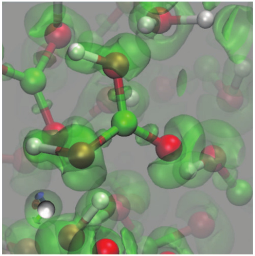
Markov chains are random sequences $(x_n)_{n\in\mathbb{N}}$ over a state space $\mathcal{X}$ defined by a transition probability

$$P(x,S) = \mathbb{P}\Big[x_{n+1} \in S \mid x_n = x\Big].$$

This is the probability of reaching the set S starting from $x \in \mathcal{X}$.

Such systems are used to model a variety of (physical) systems:

- molecules, chemical reactions;
- metals, phase transition;
- surface interactions, etc.



Typical situation

In classical mechanics, one has generally

$$x_{n+1} = x_n + F(x_n) + \sqrt{T} G_n,$$

with

- F is a force acting on the system, typically $F(x) = -\nabla H(x)$;
- the Hamiltonian $H(x)$ represents a potential energy ;
- T is a temperature, and G is a random noise.

There are natural questions about such systems:

- their long time behavior (Sec. 2.);
- the probabilities of fluctuations (Sec. 3.);
- correlations, relaxation times, linear response theory, etc.

2. Markov chains and ergodicity

Ergodicity: what and what for ?

We say that a Markov chain (x_n) has an invariant distribution μ^* when

$$x_n \sim \mu^* \Rightarrow x_{n+1} \sim \mu^*,$$

and that the process is ergodic when, for any initial distribution,

$$x_n \sim \mu^* \text{ as } n \to +\infty.$$

Some useful applications:

- more physical insight into the considered system;
- MCMC techniques;
- stability of numerical schemes;
- stochastic partial differential equations...

An ergodic theorem

M. Hairer & J. Mattingly

Assume there exist $W : \mathcal{X} \to \mathbb{R}_+$, $\gamma \in (0,1)$ and $C > 0$ such that

$$(L) \quad PW \leqslant \gamma W + C,$$

and $\alpha > 0$, $\eta \in \mathcal{P}(\mathcal{X})$ such that

$$(M) \quad \inf_{x \in \mathcal{C}} P(x,\cdot) \geqslant \alpha \eta(\cdot),$$

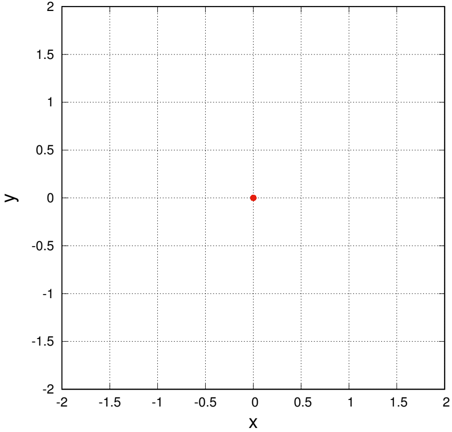
for $\mathcal{C}$ a large enough level set of W. Then, there exist a unique $\mu^* \in \mathcal{P}(\mathcal{X})$, $c > 0$ and $\lambda \in (0,1)$ such that for any $\mu \in \mathcal{P}(\mathcal{X})$

$$\|P^n \mu - \mu^*\|_W \leqslant c \lambda^n \|\mu - \mu^*\|_W.$$

Here: $\displaystyle \|f\| = \sup_{x \in \mathcal{X}} \frac{|f(x)|}{1 + W(x)}, \quad \|\mu - \nu\|_W = \sup_{\|\varphi\| \leqslant 1} \int_{\mathcal{X}} \varphi(x)(\mu - \nu)(dx).$

A basic example

Dynamics over $\mathbb{R}^2$ with $F(x;y) = (-2x;-2y)$.



The force $F = -\nabla H$ derives from the energy

$$H(x,y) = x^2 + y^2.$$

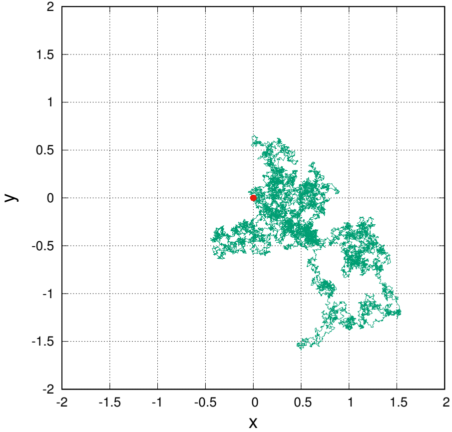
In practice, one can choose

$$W(x,y) = H(x,y),$$

the energy of the system.

A basic example

Dynamics over $\mathbb{R}^2$ with $F(x;y) = (-2x;-2y)$.



The force $F = -\nabla H$ derives from the energy

$$H(x,y) = x^2 + y^2.$$

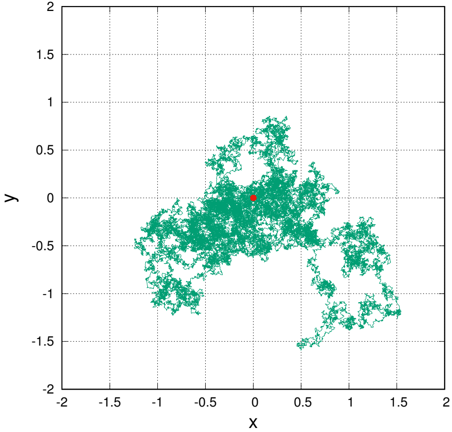
In practice, one can choose

$$W(x,y) = H(x,y),$$

the energy of the system.

A basic example

Dynamics over $\mathbb{R}^2$ with $F(x;y) = (-2x;-2y)$.



The force $F = -\nabla H$ derives from the energy

$$H(x,y) = x^2 + y^2.$$

In practice, one can choose

$$W(x,y) = H(x,y),$$

the energy of the system.

3. Feynman–Kac dynamics

Motivations

From an ergodic Markov chain (x_n), one typically computes time averages:

$$\frac{1}{n}\sum_{k=0}^{n-1} f(x_k) \xrightarrow{n\to+\infty} \int_{\mathcal{X}} f\,d\mu^*.$$

Question: what is the probability that

$$\frac{1}{n}\sum_{k=1}^{n} f(x_k) \neq \int_{\mathcal{X}} f\,d\mu^*$$

for a large but finite n?

Problem: this is a rare event, very difficult to sample numerically.

Feynman–Kac as importance sampling

Idea: give more importance to trajectories with larger values of f. In practice, replace for example

$$\mathbb{E}_x\left[\varphi(x_n)\right] \quad \text{by} \quad \mathbb{E}_x\left[\varphi(x_n)\mathrm{e}^{\sum_{k=0}^{n-1} f(x_k)}\right].$$

Problem: this semigroup does not conserve probability. In practice, we can study:

$$\Phi_n(\mu)(\varphi) = \frac{\mathbb{E}_\mu\left[\varphi(x_n)\mathrm{e}^{\sum_{k=0}^{n-1} f(x_k)}\right]}{\mathbb{E}_\mu\left[\mathrm{e}^{\sum_{k=0}^{n-1} f(x_k)}\right]},$$

or

$$\Lambda(f) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_\mu\left[\mathrm{e}^{\sum_{k=0}^{n-1} f(x_k)}\right].$$

Ergodicity for Feynman–Kac models

Define a weighted transition (assumed with some regularity):

$$P^f = \mathrm{e}^f P.$$

Ergodicity for Feynman–Kac models

Define a weighted transition (assumed with some regularity):

$$P^f = \mathrm{e}^f P.$$

Theorem [G.F., M. Rousset & G. Stoltz, in prep.]

If there exist a function $W \geqslant 1$ and sequences $\gamma_n \geqslant 0$, $b_n \geqslant 0$, compact sets K_n such that

$$(L) \quad P^f W(x) \leqslant \gamma_n W(x) + b_n \mathbb{1}_{K_n},$$

with $\gamma_n \to 0$, and that for any compact $K \subset \mathcal{X}$ there is $\eta_K \in \mathcal{P}(\mathcal{X})$, $\alpha_K > 0$ such that

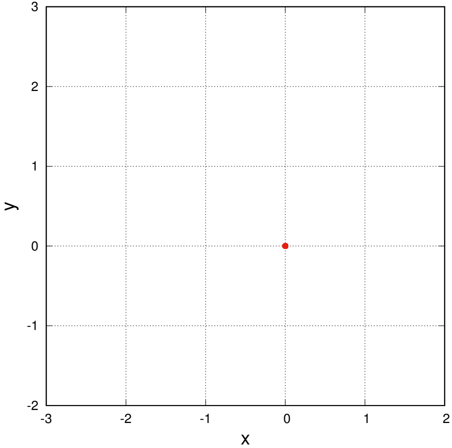
$$(M) \quad \forall\, x \in K, \quad P^f(x,\cdot) \geqslant \alpha_K \eta_K(\cdot),$$

then there exist a unique μ_f^* and $\lambda \in (0,1)$ such that for any $\mu \in \mathcal{P}(\mathcal{X})$,

$$\|\Phi_n(\mu) - \mu_f^*\|_W \leqslant C_\mu \lambda^n.$$

Numerics I: Quantum physics

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.



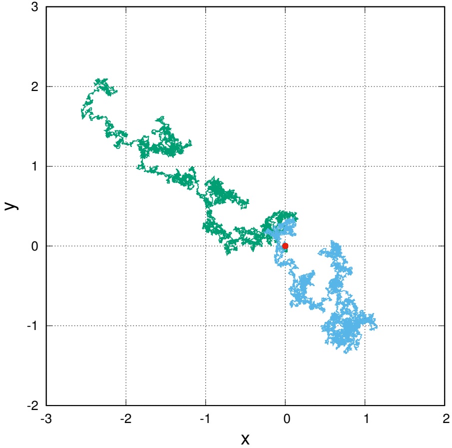
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics I: Quantum physics

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.



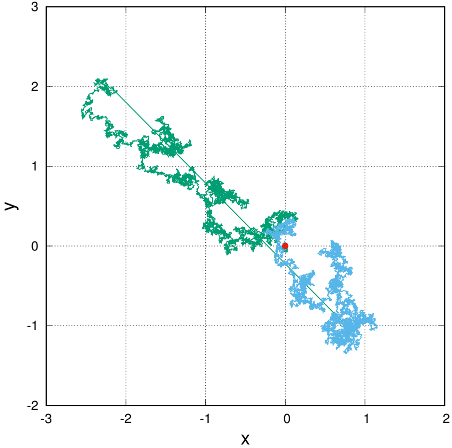
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics I: Quantum physics

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.



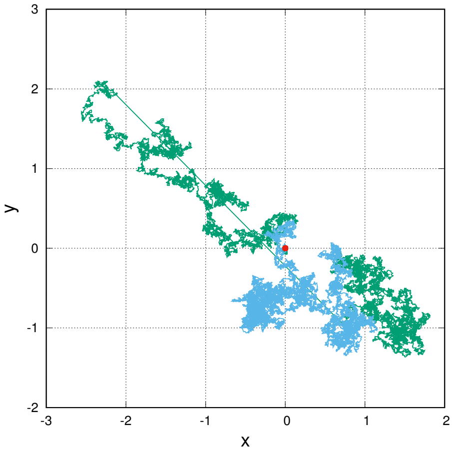
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics I: Quantum physics

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.



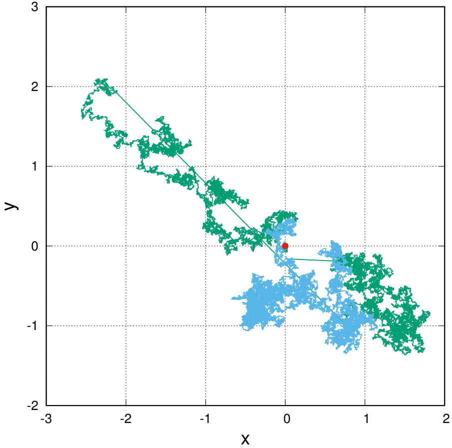
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics I: Quantum physics

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.

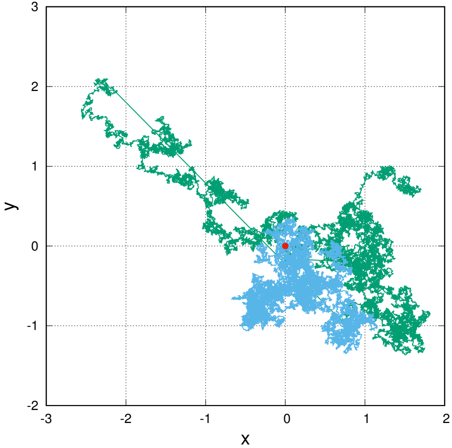


The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Brownian dynamics (x_n) over $\mathbb{R}^2$ with $f(x) = -|x|^2$.



The weight

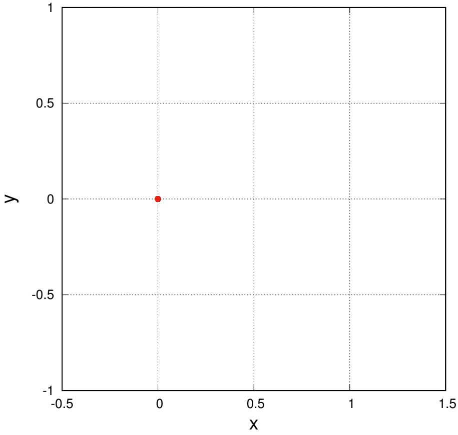
$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m) \right).$$

of particle m at time k is used for resampling (renormalization).

Conclusion: there is a confinement by selection, *i.e.* through f.

Numerics II: importance sampling

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x, -2y)$. Weight function $f(x,y) = x$.



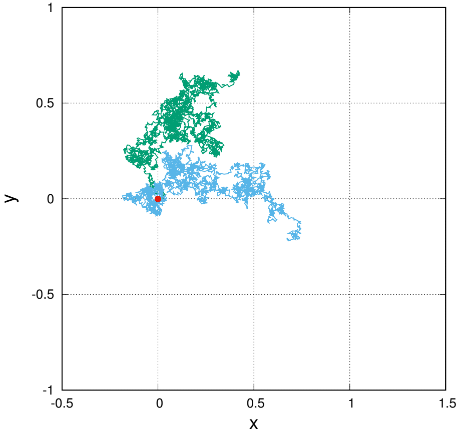
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics II: importance sampling

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x,-2y)$. Weight function $f(x,y) = x$.



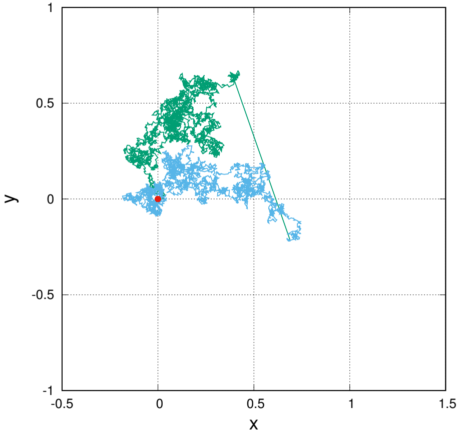
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics II: importance sampling

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x, -2y)$. Weight function $f(x,y) = x$.



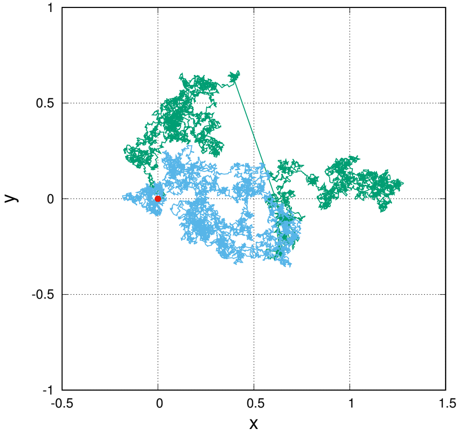
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics II: importance sampling

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x,-2y)$. Weight function $f(x,y) = x$.

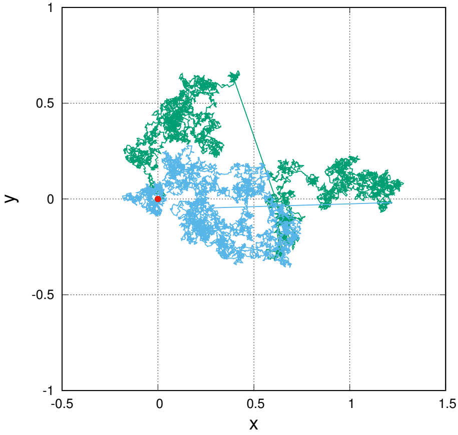


The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x, -2y)$. Weight function $f(x,y) = x$.



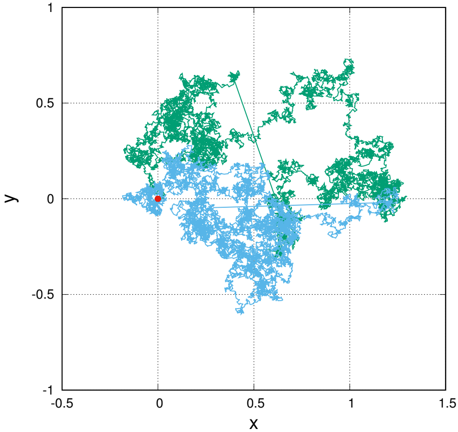
The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Numerics II: importance sampling

Ornstein-Uhlenbeck dynamics with $F(x,y) = (-2x,-2y)$. Weight function $f(x,y) = x$.



The weight

$$w_m^k = \exp\left(\sum_{i=0}^{k-1} f(x_i^m)\right).$$

of particle m at time k is used for resampling (renormalization).

Conclusion: the particles are selected towards the right.

Conclusion

Take home message:

- problems of ergodicity, long time behavior;
- Feynman–Kac dynamics: **funny problem** (as far as mathematics can be funny);
- new results, «energy» for non-probabilistic operators;
- further works: consequences for large deviations of additive functionals.