CONVERGENCE IN DISTRIBUTION OF SOME SELF-INTERACTING DIFFUSIONS

ALINE KURTZMANN

Abstract. The present paper is concerned with some self-interacting diffusions $(X_t, t \geq 0)$ living on $\mathbb{R}^d$. These diffusions are solutions to stochastic differential equations:

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t$$

where $\overline{\mu}_t$ is the empirical mean of the process X, V is an asymptotically strictly convex potential and g is a given positive function. We study the asymptotic behaviour of X for three different families of functions g. If $g(t) = k \log t$ with k small enough, then the process X converges in distribution towards the global minima of V, whereas if $tg(t) \to c \in]0, +\infty]$ or if $g(t) \to g(\infty) \in [0, +\infty[$, then X converges in distribution iif $\int xe^{-2V(x)}\,\mathrm{d}x = 0$.

1. Introduction

The aim of this paper is to obtain necessary and sufficient conditions for the convergence in distribution of a self-interacting diffusion living on $\mathbb{R}^d$. Consider a smooth potential $V : \mathbb{R}^d \to \mathbb{R}_+$ and a map $g : \mathbb{R}_+ \to \mathbb{R}_+$. We study the asymptotic behaviour of the self-interacting diffusion X given by

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t, \quad X_0 = x = \bar{\mu}_0 \tag{1.1}$$

where B is a standard Brownian motion and $\overline{\mu}_t$ denotes the empirical mean of the process X:

$$\overline{\mu}_t = \frac{1}{t}\int_0^t X_s \mathrm{d}s. \tag{1.2}$$

This is a model of reinforcement that could be used to represent the (simplified) behaviour of some social insects. Some insects, as ants, mark their paths with pheromones. This serves as a guide for other ants to return to the nest. The trail of pheromones is denoted by X and its evaporation by g. Despite this evaporation, the path is reinforced and the insects gradually manage to find the best route.

The same model has been already studied by Chambeu & Kurtzmann [4], in case of an unbounded increasing function g. The authors have proven that, under certain conditions, the process satisfies a kind of pointwise ergodic theorem, and that if V admits a unique minimum at 0, then X_t converges almost surely. In this paper, we do not suppose that g increases to the infinity nor that V admits a unique minimum at 0. This will obviously change the asymptotic behaviour of X, even if X will converge in distribution in most of the cases. We will essentially use two different techniques here. The first one is the well-known theory of simulated annealing, which has been developed a lot since the 80's with a huge literature, whereas the second one is simply a change of scale added to a change of "speed measure".

Let us explain briefly the simulated annealing method. An important question for physical systems is to find the globally minimum energy states of the system. Experimentally, the ground states are reached by chemical annealing. One first melts a substance and then cools

Date: February 18, 2014.

it slowly, being careful to pass slowly through the freezing temperature. If the temperature decreases too rapidly, then the system does not end up in a ground state, but in a local non-global minimum. On the other hand, if the temperature decreases too slowly, then the system approaches the ground states very slowly. The competition between these two effects determines the optimal speed of cooling, that is the annealing schedule. The study of simulated annealing has involved the theory of non-homogeneous Markov chains and diffusion processes, large deviation theory, spectral analysis of operators and singular perturbation theory. Pioneering work was done by Freidlin and Wentzell [6]. The initial problem consists in finding the global minima of a given function U. Actually, one has to study the diffusion Markov process X^ε in $\mathbb{R}^d$ given by the Langevin-type Markov diffusion $\mathrm{d}X_t^\varepsilon = \mathrm{d}B_{\varepsilon_t} - \nabla U(X^\varepsilon)\mathrm{d}t$. If the temperature ε is constant for a sufficiently large amount of time, then the process X^ε and the fixed temperature process behave approximatively the same at the end of that time interval. The optimal annealing schedule, that is ε for the convergence criterion $\mathbb{P}(X_t^\varepsilon \in Min) \underset{t\to\infty}{\longrightarrow} 1$, where Min denotes the set of all the global minima of U, was first determined by Hajek [8] for a finite state space. Chiang, Hwang and Sheu [5] studied the convergence rate of $\mathbb{P}_x(X_t^\varepsilon \in \cdot)$ via the large deviations of the transition density of X^ε. They were one of the first to show the convergence of the algorithm of the simulated annealing for $\varepsilon_t^2 = k/\log t$, for k large enough, related to the second eigenvalue of the corresponding (to X^ε) infinitesimal generator. Finally, Holley and Stroock [11] initiated a new method and proved, in the discrete case, the convergence of the simulated annealing algorithm via the Sobolev inequality. They went further in their study with Kusuoka [9]. Later, Miclo [15] proved, through some functional inequalities, that the free energy (that is the relative entropy of the distribution of the process at time t with respect to the invariant probability measure for the elliptic operator considered as a time-homogeneous operator by fixing t) satisfies a differential inequality, which implies (under some decreasing evolution of the temperature to zero) the convergence of the process to the global minima of the potential. And if the temperature ε decreases too fast to zero, then the potential can freeze in a local minimum (depending on the initial condition) and so the process converges to this local minimum.

We begin to study the $\mathbb{R}^d$-valued Markov process $Y_t := X_t - \bar{\mu}_t$, which satisfies the following SDE

$$\begin{cases} \mathrm{d}Y_t = \mathrm{d}B_t - g(t)\nabla V(Y_t)\mathrm{d}t - Y_t\frac{\mathrm{d}t}{t}, \quad Y_0 = 0; \\ \mathrm{d}\overline{\mu}_t = Y_t\frac{\mathrm{d}t}{t}, \quad \overline{\mu}_0 = x. \end{cases} \tag{1.3}$$

We will adapt the simulated annealing method to Y for functions g large enough (that is g does not go to zero) to prove the convergence in distribution of Y.

We wish to point out that a one-dimensional Brownian motion in a time-dependent potential has been recently studied by Gradinaru and Offret [7]:

$$\mathrm{d}Z_t = \mathrm{d}B_t - \partial_x V_{\rho,\alpha,\beta}(t, Z_t)\mathrm{d}t, \quad Z_{t_0} = z_0$$

with $V_{\rho,\alpha,\beta}(t,x) = \frac{\rho}{\alpha+1}\frac{|x|^{\alpha+1}}{t^\beta}\mathbb{1}_{\alpha\neq-1} + \rho\frac{\log|x|}{t^\beta}\mathbb{1}_{\alpha=-1}$ and $z_0, \rho, \alpha, \beta \in \mathbb{R}$. This is quite close in spirit to the study of our process Y, even if the authors suppose in [7] that both V and $1/g$ are polynomial. They obtain conditions for the recurrence, transience and convergence of the studied process Z. We refer to the survey of Ivanov et al [13] for the existence and uniqueness of solutions to such equations. In the present paper, we do not suppose that g is polynomial and the dimension is $d \geq 1$, and thus we obtain less precise results.

The remainder of the paper is organized as follows. First, in Section 2, we introduce some useful tools, such as the logarithmic Sobolev inequality and the Kullback information. They both will be needed for the simulated annealing study. Section 3 is devoted to the simulated annealing method in the case when g behaves asymptotically as $k \log t$. In this part, we will prove the (pointwise) ergodicity of the process Y and the convergence in distribution of X, depending on the potential V. Finally, Section 4 deals with the convergence in distribution of X when $tg(t) \to +\infty$ and $g(t) \to c \geq 0$, depending on the asymptotics of V.

2. Some useful tools

2.1. **Assumptions and existence.** In the whole following, $(\cdot, \cdot)$ denotes the Euclidiean scalar product. We denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$. We denote by G the function $G(t) = \int_0^t g(s)\mathrm{d}s$. We assume that the mapping $g : \mathbb{R}_+ \to \mathbb{R}_+$ is $\mathcal{C}^1(\mathbb{R}_+)$. The precise hypothesis on g will be given at the beginning of each section.

In the sequel, the technical assumptions on the potential $V : \mathbb{R}^d \to \mathbb{R}_+$ are the following:

(1) (*regularity and positivity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $V \geq 0$;
(2) (*convexity*) $V = W + \chi$ where W is $C_W(>0)$-strictly uniformly convex and χ is a compactly supported function and there exists $C_\chi > 0$ such that $\nabla\chi$ is C_χ-Lipschitz ;
(3) (*growth*) there exists $a > 0$ such that for all $x \in \mathbb{R}^d$, we have

$$\Delta V(x) \leq aV(x) \quad \text{and} \quad \lim_{|x|\to+\infty} \frac{|\nabla V(x)|^2}{V(x)} = +\infty. \tag{2.1}$$

We also assume that V has a finite number of critical points. Let $Max = \{M_1, M_2, \cdots, M_p\}$ be the set of the saddle points and local maxima of V and $Min = \{m_1, m_2, \cdots, m_n\}$ be the set of the local minima of V, such that the Hessian matrix is non-degenerate for all local minimum. Without any loss of generality, we suppose that $\min V = 0$.

Remark 2.1. *The case V of quadratic growth is excluded here, as it has been fully studied in* [4].

Let us first prove the global strong existence and uniqueness of the process X.

Proposition 2.1. *For any $x \in \mathbb{R}^d, \mu \in \mathcal{P}(\mathbb{R}^d)$, there exists a unique global strong solution $(X_t, t \geq 0)$ of* (1.1).

Proof. The local existence and uniqueness of such a process is standard. We only need to prove here that Y, hence X (because $X_t := x + Y_t + \int_0^t Y_s \frac{\mathrm{d}s}{s}$), does not explode in a finite time. To this aim, apply Itô's formula to the function $x \mapsto V(x)$:

$$\mathrm{d}V(Y_t) = (\nabla V(Y_t), \mathrm{d}B_t) + \left(\frac{1}{2}\Delta V(Y_t) - g(t)|\nabla V(Y_t)|^2 - \frac{1}{t}(\nabla V(Y_t), Y_t)\right)\mathrm{d}t,$$

and introduce the sequence of stopping times $\tau_0 = 0$ and

$$\tau_n = \inf\{t \geq 0; V(Y_t) > n\}.$$

By the convexity condition, we have $(\nabla V(y), y) \underset{|y|\to+\infty}{\longrightarrow} +\infty$, and by the condition (2.1), there exists $C > 0$ such that $\mathbb{E}V(Y_{t\wedge\tau_n}) \leq \mathbb{E}V(Y_0) + Ct$. □

2.2. **Preliminaries.**

2.2.1. *Logarithmic Sobolev inequality.*

Definition 2.2. *The probability measure μ satisfies the logarithmic Sobolev inequality, with the constant C_{LS}, denoted by $LSI(C_{LS})$, if for all function $h \in L^2(\mu)$, we have*

$$\int h^2 \log h^2 \mathrm{d}\mu - \left(\int h^2 \mathrm{d}\mu\right) \log\left(\int h^2 \mathrm{d}\mu\right) \le C_{LS} \int |\nabla h|^2 \mathrm{d}\mu.$$

Let $p(s,x,t,y)$ denote the density of the semi-group corresponding to the non-homogeneous Markov process Z, defined by

$$\mathrm{d}Z_t = \varepsilon_t \mathrm{d}B_t - \left(\nabla V(Z_t) + \frac{Z_t}{a_t}\right)\mathrm{d}t.$$

We will specify later the precise form of ε_t and a_t. We associate to this process the probability measure $\Pi_{t,\varepsilon_t}(\mathrm{d}x) = \frac{1}{\pi_t}\exp\{-2\varepsilon_t^{-2}(V(x)+|x|^2/2a_t)\}\mathrm{d}x$, where π_t is the normalization constant of Π_{t,ε_t}.

Lemma 2.3. *The family of probability measures $(\Pi_{t,\varepsilon_t}, t \ge 0)$ satisfies a logarithmic Sobolev inequality $LSI(C(t))$.*

Proof. We use the celebrated Bakry-Emery Γ_2-criterion, see [1]. We recall that, to the operator L_{t,ε_t}, we associate the operator "carré du champ", that is (for all function $f, g \in \mathcal{C}^\infty$)

$$\Gamma_t^V(f,g) := \frac{1}{2}\left(L_{t,\varepsilon_t}(fg) - fL_{t,\varepsilon_t}g - gL_{t,\varepsilon_t}f\right). \tag{2.2}$$

Then, we define the operator Γ_2^V as

$$\Gamma_2^V(t)(f) := \frac{1}{2}\left(L_{t,\varepsilon_t}\Gamma_t^V(f,f) - 2\Gamma_t^V(f, L_{t,\varepsilon_t}f)\right). \tag{2.3}$$

The Γ_2-criterion asserts that if there exists a positive constant C such that $\Gamma_2^{V_t} \ge C\Gamma_t^{V_t}$, then Π_{t,ε_t} satisfies a logarithmic Sobolev inequality, with the constant $2/C$.

An easy calculation, for any function f of class $\mathcal{C}^\infty$, leads to

$$\Gamma_t^V(f,f) = \varepsilon_t^2|\nabla f|^2$$

and

$$\Gamma_2^V(t)(f) = \frac{\varepsilon_t^2}{2}(\nabla f, \nabla^2 V \nabla f) + \frac{\varepsilon_t^4}{4}||\nabla^2 f||^2 + \frac{\varepsilon_t^2}{2a(t)}|\nabla f|^2.$$

As V (and also V_t) is strictly convex off a compact set, we have the decomposition $V = W+\chi$ as in the convexity hypothesis. We apply the Γ_2-criterion of Bakry-Emery to the function W and we get that $\Gamma_2^W(t)(f) \ge C_W \Gamma_t^W(f)$. Thus, the probability measure $e^{-2\varepsilon_t^{-2}(W(x)+|x|^2/a(t))}/\pi_t$ satisfies the inequality $LSI(2/C_W)$. We conclude, by the perturbation lemma due to Holley and Stroock [10], that the measure Π_{t,ε_t} satisfies a Sobolev logarithmic inequality with a constant less than or equal to $2e^{\frac{2}{\varepsilon_t^2} osc\chi}/C_W$, where $osc(\chi) = \sup\chi - \inf\chi$. □

2.2.2. *Kullback information.*

Definition 2.4. *We define the free energy (up to an additive constant), known as the relative Kullback information, of a probability measure ν with respect to a probability measure Π by:*

$$H(\nu|\Pi) := \int \mathrm{d}\nu \log\frac{\mathrm{d}\nu}{\mathrm{d}\Pi}.$$

If we suppose that ν (respectively Π) has the density ν (respectively π) with respect to the Lebesgue measure λ, then one has

$$H(\nu|\Pi) := \int \nu \log \frac{\nu}{\pi} \mathrm{d}\lambda. \tag{2.4}$$

In this paper, we will first prove the decrease to zero of the relative free energy of the law of Z_t with respect to Π_{t,ε_t}. The classical Csiszár-Kullback-Pinsker inequality relates the total variation norm to the free energy in the following way (see for instance [11]):

$$\|\mu - \nu\|_{TV} \leq \sqrt{2H(\mu|\nu)}. \tag{2.5}$$

So, as the total variation norm metrizes the convergence in distribution, once we have proven that the measure Π_{t,ε_t} converges weakly to a measure Π and $H(p_t|\Pi_{t,\varepsilon_t})$ goes to zero, then the distribution of Z_t converges to Π. As Z_t is the time-shifted process Y_t, we obtain this way that Y converges in distribution to Π.

Our strategy to show that $H(p_t|H_{t,\varepsilon_t})$ goes to zero is the following. To shorten notation, let $p_t := p(t_0, x_0, t, \cdot)$ be the distribution law of the process Z_t conditioned on $Z_{t_0} = x_0$. We recall that the family of probability measures $(\Pi_{t,\varepsilon_t}, t \geq 0)$ satisfies a Sobolev logarithmic inequality $LSI(C(t))$. We have also $\Pi_{t,\varepsilon_t}(\mathrm{d}x) = \pi_{t,\varepsilon_t}(x)\lambda(\mathrm{d}x)$. So, we choose $h_t = \sqrt{\frac{p_t}{\pi_{t,\varepsilon_t}}}$ satisfying $\int h_t^2 \mathrm{d}\Pi_{t,\varepsilon_t} = 1$ and we will show in Corollary 3.6 the existence of $C(t) > 0$ such that

$$H(p_t|\Pi_{t,\varepsilon_t}) = \int p_t \log \frac{p_t}{\pi_{t,\varepsilon_t}} \mathrm{d}\lambda \leq C(t) \int |\nabla h_t|^2 \mathrm{d}\Pi_{t,\varepsilon_t}. \tag{2.6}$$

2.2.3. *Asymptotic pseudotrajectories.* In Section 4, we will use the notion of asymptotic pseudotrajectory, introduced by Benaïm and Hirsch [2]. It is particularly useful to analyze the long-term behaviour of stochastic processes, considered as approximations of solutions of ordinary differential equation (the "ODE method").

Definition 2.5. *The process Y is an asymptotic pseudotrajectory for the flow ϕ if $\forall T > 0$*

$$\lim_{t \to +\infty} \sup_{0 \leq s \leq T} |Y_{t+s} - \phi_s(Y_t)| = 0 \ a.s. \tag{2.7}$$

It is shown in [2] that if Y is an asymptotic pseudotrajectory for ϕ, then the ω-limit set of the flow generated by ϕ is the same as the ω-limit set of the process Y.

3. The simulated annealing method

Assume that the mapping $g : \mathbb{R}_+ \to \mathbb{R}_+$ is asymptotically equivalent (up to a multiplicative positive constant) to $\log t$ and satisfies $g \in \mathcal{C}^1(\mathbb{R}_+)$ is such that $g(0) > 0$ and for all $T > 0$, $G^{-1}(t+T) - G^{-1}(t) \underset{t\to\infty}{\longrightarrow} 0$ where G^{-1} is the generalized inverse of G.

Instead of considering Y, we consider the time-changed process $Z_t := Y_{G^{-1}(t)}$. This last process satisfies the following SDE

$$\mathrm{d}Z_t = \frac{1}{\sqrt{g \circ G^{-1}(t)}} \mathrm{d}B_t - \left(\nabla V(Z_t) + \frac{Z_t}{G^{-1}(t) g \circ G^{-1}(t)} \right) \mathrm{d}t, \ \ Z_0 = Y_{G^{-1}(0)} = 0, \tag{3.1}$$

where B is a Brownian motion such that $\int_0^t \frac{1}{\sqrt{g \circ G^{-1}(s)}} \mathrm{d}B_s$ has the same law as $B_{G^{-1}(t)}$.

3.1. **Convergence in distribution towards the global minima of V.** We define $\varepsilon_t^2 = \frac{1}{g\circ G^{-1}(t)}$ and $a(t) = G^{-1}(t) g \circ G^{-1}(t)$. The process Z satisfies

$$\mathrm{d}Z_t = \varepsilon_t \mathrm{d}B_t - \nabla V_t(Z_t)\mathrm{d}t \tag{3.2}$$

where we have defined $V_t(x) := V(x) + \frac{|x|^2}{2a(t)}$. Actually, we will prove that this non-homogeneous Markov process converges in distribution to a measure that could correspond to its "invariant" probability measure. Of course, if we suppose that $a(t) \equiv a$ and $\varepsilon_t \equiv \varepsilon$, then the convergence in distribution is obvious. It happens that the spectral gap λ appears naturally in our study. Heuristically, when the time is of order $e^{\varepsilon^{-2}\lambda}$, the process is very close to the probability measure

$$\Pi_{t,\varepsilon}(\mathrm{d}x) := \frac{1}{\pi(t,\varepsilon)} e^{-2\varepsilon^{-2}V_t(x)}\mathrm{d}x. \tag{3.3}$$

It remains to show the convergence of $\Pi_{t,\varepsilon}$ when t goes to the infinity.

Let $L_{t,\varepsilon}$ be the operator defined by $L_{t,\varepsilon} := \frac{1}{2}\varepsilon^2\Delta - (\nabla V_t, \nabla)$. As $|\nabla V_t|^2 - \Delta V_t$ goes to the infinity as $|x| \to \infty$, the theory of Schrödinger operator (see for instance [16, Thm13.6]) implies that $L_{t,\varepsilon}$ is self-adjoint in $L^2(\Pi_{t,\varepsilon})$ and the spectrum of $L_{t,\varepsilon}$ is discrete: $0 = \lambda_1(t,\varepsilon) < -\lambda_2(t,\varepsilon) < \ldots$. The subspace corresponding to the first eigenvalue $\lambda_1(t,\varepsilon)$ is composed of the constant functions and so

$$\lambda_2(t,\varepsilon) = \inf\left\{\int |\nabla\phi|^2 \mathrm{d}\Pi_{t,\varepsilon}; \quad \mathrm{Var}_{\Pi_{t,\varepsilon}}(\phi) = 1, \phi \in \mathcal{D}(\mathbb{R}^d)\right\}.$$

Our first aim is to compute the eigenvalue λ_2 and study its behaviour when $t \to \infty$.

Lemma 3.1. *Let $\varepsilon > 0$ be fixed. The probability measure $\Pi_{t,\varepsilon}$ converges weakly, as $t \to \infty$, to $\Pi_{\infty,\varepsilon}(\mathrm{d}x) := \frac{1}{\pi(\varepsilon)} e^{-2\varepsilon^{-2}V(x)}\mathrm{d}x$. Moreover, $\lim_{\varepsilon\to 0}\Pi_{\infty,\varepsilon}$ exists and is denoted by Π_0.*

Proof. We only need to recall that $\varepsilon_t^2 a(t) = G^{-1}(t)$ diverges with t. More explicitly, the normalization constant is

$$\pi(t,\varepsilon) = \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2}}\mathrm{d}x.$$

Let K be the compact set $K := \{x | V(x) \leq 1\}$. There exists a constant $A > 0$ such that K is included in the ball centered in 0 and with radius A. Then, on one hand, we get

$$\int_{K^c} e^{-2\varepsilon^{-2}V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2}}\mathrm{d}x \leq \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}} e^{-2\frac{|x|^2}{a(t)\varepsilon^2}}\mathrm{d}x \leq C(a(t)\varepsilon^2)^{d/2} e^{-2\varepsilon^{-2}}.$$

On the other hand we obtain,

$$\int_K e^{-2\varepsilon^{-2}V(x)}\mathrm{d}x \geq \int_K e^{-2\varepsilon^{-2}V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2}}\mathrm{d}x \geq \int_K e^{-2\varepsilon^{-2}V(x)} e^{-2\frac{A^2}{a(t)\varepsilon^2}}\mathrm{d}x.$$

But we know by the Laplace formula (see [12]) that

$$\int_K e^{-2\varepsilon^{-2}V(x)}\mathrm{d}x \underset{t\to+\infty}{\sim} \sum_i (2\pi\varepsilon^2)^{d/2} (\det \nabla^2 V(x_i))^{-1/2}$$

where $(x_i)_i$ are the global minima of V (we recall that they form a finite set). As a consequence,

$$\pi(t,\varepsilon) \underset{t\to+\infty}{\sim} \sum_i (2\pi\varepsilon^2)^{d/2} (\det \nabla^2 V(x_i))^{-1/2}.$$

By the same method, if ϕ is a continuous function with compact support containing for example only the global minimum x_1, we have

$$\int_{\mathbb{R}^d} \phi(x) e^{-2\varepsilon^{-2}V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2}} \mathrm{d}x \underset{t\to+\infty}{\sim} (2\pi\varepsilon^2)^{d/2} (\det \nabla^2 V(x_1))^{-1/2} \phi(x_1).$$

This gives the explicit form of $\lim_{\varepsilon\to 0} \Pi_{\infty,\varepsilon}(\mathrm{d}x) = \Pi_0(\mathrm{d}x)$. □

Consider for a moment $\Pi_{\infty,\varepsilon}$. We remark that V_t converges to V when t goes to infinity. Hwang established in [12] that $\Pi_{\infty,\varepsilon}$ converges weakly when ε converges to zero. Let N be the set of the global minima of V. Hwang has proved the following:

- if $\lambda(N) > 0$ (where λ is the Lebesgue measure on $\mathbb{R}^d$), then $\Pi_{\infty,\varepsilon}$ converges weakly to $\frac{1}{\lambda(N)} \mathbb{1}_N \mathrm{d}x$;
- if $N = \{x_1, \ldots, x_n\}$ then $\Pi_{\infty,\varepsilon}$ converges weakly to
$$\frac{1}{\sum_{1\le i\le n} (\det \nabla^2 V(x_i))^{-1/2}} \sum_{1\le i\le n} (\det \nabla^2 V(x_i))^{-1/2} \delta_{x_i};$$
- more generally, suppose that N is the finite union of some smooth manifolds ($\mathcal{C}^3$), and each component is a compact connected smooth manifold and the determinant of the Hessian (normal to N in $x \in N$) $\det(\nabla^2 V(x))$ is not identically zero. Then, there exists a probability measure $\mathcal{M}$, on the highest dimensional manifolds, such that $\Pi_{\infty,\varepsilon}$ converges weakly to $\frac{1}{\int (\det \nabla^2 V(x))^{-1/2} \mathcal{M}(\mathrm{d}x)} (\det \nabla^2 V(x))^{-1/2} \mathcal{M}(\mathrm{d}x)$.

We adapt to our setting the results of Hwang in the following proposition.

Proposition 3.2. *The probability measure Π_{t,ε_t} converges weakly to Π_0 as t goes to infinity. Moreover, the probability measure Π_0 concentrates on the global minima of V.*

Proof. The result of Hwang shows that the probability measure $\Pi_{\infty,\varepsilon_t}$ converges weakly to Π_0 as t goes to the infinity, and the probability measure Π_0 concentrates on the global minima of V. We combine this result with Lemma 3.1 to prove the proposition. □

In order to show that Z converges in distribution to a measure supported on the global minima of V, we need two more technical results. We mix the approaches initiated by Holley, Kusuoka & Stroock [9] and Miclo [15]. Indeed, we will use some functional inequalities, and show that the free energy (corresponding to our process) decreases. We suppose in the following that $g \circ G^{-1}(t) = \frac{\log t}{k}$ for some k sufficiently large (and the same proof actually reads when $g \circ G^{-1}(t)$ is asymptotically equivalent to $\frac{\log t}{k}$).

Definition 3.3. *The maximal height of the function V_t is the non-negative function $m(t)$ defined by*

$$m(t) := \sup\{H_t(x) - V_t(x); \quad x \in K\}, \tag{3.4}$$

where

$$\begin{aligned} H_t(x) &:= \inf\{E_t(\gamma); \quad \gamma \in \mathcal{C}^1([0,1], K); \gamma(0) = x, \gamma(1) = 0\}, \\ E_t(\gamma) &:= \sup\{V_t(\gamma(u)); \quad u \in [0,1]\}. \end{aligned}$$

Remark 3.1. *1) The function $m(t)$ corresponds to the maximum of all the minimal energies needed to go from each point of $\mathbb{R}^d$ to 0.*
2) The function $m(t)$ is positive if and only if there exist more than one local minimum of V.

Lemma 3.4. *We have that* $\lim_{t\to\infty} m(t) = m$*, where* m *is the maximal height function corresponding to* V*.*

Proof. Let $M := \sup\{|x|^2; V(x) \leq 1\}$. For any path γ, we easily have $E_t(\gamma) \leq E_\infty(\gamma) + \frac{M}{a(t)}$. Then, by definition of H_t, we get

$$|H_t(x) - H_\infty(x)| \leq \frac{M}{a(t)}.$$

As a consequence, there exists $C > 0$ such that

$$|m(t) - m| \leq \sup\left\{\left|H_t(x) - H_\infty(x) - \frac{|x|^2}{a(t)}\right|; \quad x \in K\right\} \leq \frac{C}{a(t)}$$

and the result follows. □

A very important theorem permits one to relate the height function to the second eigenvalue of the infinitesimal generator of Y^ε (that is the constant involved in the spectral gap inequality).

Theorem 3.5. *(Jacquot [14], Thm 1.1) The invariant measure* $\Pi_{t,\varepsilon}$ *admits a spectral gap* λ_2*: there exist* $C_1, C_2, \varepsilon_0 > 0$ *such that for all* $\varepsilon > \varepsilon_0$*, one has for all continuous* $f \in L^2(\Pi_{t,\varepsilon})$

$$||P_s^{t,\varepsilon} f - \Pi_{t,\varepsilon} f||_{L^2(\Pi_{t,\varepsilon})} \leq e^{-\left(2\varepsilon^{-2}m(t)+\log Q(\varepsilon)-\log(2-\varepsilon^2)\right)s}\mathrm{Var}_{\Pi_{t,\varepsilon}}(f) = e^{-2\lambda_2(t,\varepsilon)}\mathrm{Var}_{\Pi_{t,\varepsilon}}(f),$$

where $Q(\varepsilon) = C_W\varepsilon^2 + C_1\varepsilon^{-6d}(1+C_2\varepsilon^{-2d+2}) + \frac{\varepsilon^2}{C_W^{-1}-(d-1)\varepsilon^2}$*. Moreover,* $\lim_{\varepsilon\to 0}\varepsilon^2 \log \lambda_2(\infty, \varepsilon) = -2m$*.*

Corollary 3.6. *The family of probability measures* $(\Pi_{t,\varepsilon_t}, t \geq 0)$ *satisfies a logarithmic Sobolev inequality* $LSI(C(t))$*, with* $C(t) = Q(\varepsilon_t)e^{2\varepsilon_t^{-2}m(t)}$*.*

Proof. Hölder's inequality implies that the logarithmic Sobolev constant is smaller than the inverse of the spectral gap constant in Theorem 3.5. □

We will now use some functional inequalities in order to prove the convergence of Z_t (and thus Y_t) towards the global minima of V. Let $p(s, x, t, y)$ denote the density of the semigroup corresponding to the non-homogeneous Markov process Z.

Theorem 3.7. *Suppose that* $\varepsilon_t^2 = k/\log t$*, where* $k > 2m$*. Then, for all initial* t_0, x_0*, the free energy* $H\left(p(t_0, x_0, t, \cdot)|\Pi_{t,\varepsilon_t}\right)$ *converges to 0 as* t *goes to the infinity.*

To prove Theorem 3.7, we need the three following technical results. We will first state them all, postponing there proofs, and deduce from them the latter Theorem 3.7. Let us state the first technical result.

Proposition 3.8. *For all initial* t_0, x_0*, we get*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H\left(p(t_0,x_0,t,\cdot)|\Pi_{t,\varepsilon_t}\right) &\leq -\frac{2}{C(t)}\varepsilon_t^2 H\left(p(t_0,x_0,t,\cdot)|\Pi_{t,\varepsilon_t}\right) - 4\dot{\varepsilon}(t)\varepsilon_t^{-3}\int p(t_0,x_0,t,\cdot)(V_t - <V_t>_{\Pi_{t,\varepsilon_t}})\mathrm{d}\lambda \\ &+ \frac{2}{\varepsilon_t^2}\int p(t_0,x_0,t,\cdot)(\dot{V}_t - <\dot{V}_t>_{\Pi_{t,\varepsilon_t}})\mathrm{d}\lambda,\end{aligned}$$

where we have denoted $\dot{V}_t = \frac{\partial}{\partial t}V_t$*.*

Lemma 3.9. *(Miclo, Lemma 6) Let $f : [0,\infty[\to \mathbb{R}_+$ be a continuous function such that a.s.*

$$f'(t) \leq \alpha_t - \beta_t f(t),$$

where α and β are two continuous non-negative functions such that $\int_1^\infty \beta_t \mathrm{d}t = \infty$ and $\lim_{t\to\infty} \alpha_t/\beta_t = 0$. Then $\lim_{t\to\infty} f(t) = 0$.

We now need a technical lemma to conclude that the free energy converges to 0.

Lemma 3.10. *For all $t \geq 0$, the quantity $< |x|^2 >_{\Pi_{t,\varepsilon_t}}$ is bounded.*

We are now ready to prove Theorem 3.7.

Proof of Theorem 3.7. Let $t_0 \geq 0$ and $x_0 \in \mathbb{R}^d$. Consider the process Z_t, solution to the SDE

$$\mathrm{d}Z_t = \varepsilon_t \mathrm{d}W_t - \left(\nabla V(Z_t) + \frac{Z_t}{a(t)}\right)\mathrm{d}t, \;\; Z_{t_0} = x_0.$$

We can rewrite the result of Proposition 3.8 in the following way, where we remind that $p_t = p(t_0, x_0, t, \cdot)$ denotes the distribution law of the process Z conditioned on $Z_{t_0} = x_0$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi_{t,\varepsilon_t}) \;\; &\leq \;\; -\frac{2}{C(t)}\varepsilon_t^2 H(p_t|\Pi_{t,\varepsilon_t}) + \frac{2}{\varepsilon_t^2}(\mathbb{E}\dot{V}_t(Z_t) - <\dot{V}_t>_{\Pi_{t,\varepsilon_t}}) \\ &- \;\; 4\dot{\varepsilon}_t\varepsilon_t^{-3}\left(\mathbb{E}V_t(Z_t) - <V_t>_{\Pi_{t,\varepsilon_t}}\right).\end{aligned}$$

We remind to the reader that $V(x) \geq c|x|^2$ out of a compact set and it is proved in [4] that $\mathbb{E}V(Z_t) = O(1)$. We therefore have $\mathbb{E}V_t(Z_t) = O(1)$. Moreover, the function $t \mapsto a(t)$ is non-decreasing while $t \mapsto \varepsilon_t$ is non-increasing. Thus, as $\dot{V}_t(x) = -\frac{\dot{a}(t)}{2a(t)^2}|x|^2$, the two terms $\mathbb{E}(\dot{V}_t(Z_t))$ and $<V_t>_{\Pi_{t,\varepsilon_t}}$ do not play any role in the upper bound. It now remains to find a upper bound for $<\dot{V}_t>_{\Pi_{t,\varepsilon_t}}$. To this aim, we use Lemma 3.10. Indeed, there exist two positive constants M_1, M_2 such that

$$\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi_{t,\varepsilon_t}) \leq -\frac{2}{C(t)}\varepsilon_t^2 H(p_t|\Pi_{t,\varepsilon_t}) - M_1\frac{\dot{\varepsilon}(t)}{\varepsilon_t^3} + M_2\frac{\dot{a}(t)}{\varepsilon_t^2 a(t)^2}.$$

We now use Lemma 3.9. We easily compute the time-derivative of $a(t)$:

$$\frac{\dot{a}(t)}{a(t)^2\varepsilon_t^4} = -\frac{\dot{\varepsilon}(t)}{\varepsilon_t^3 G^{-1}(t)} + \frac{1}{(g\circ G^{-1}(t))G^{-1}(t)^2\varepsilon_t^2}.$$

Using the explicit expression of ε_t, that is $\varepsilon_t^2 = k/\log t$, we have

$$C(t)\frac{\dot{a}(t)}{a(t)^2\varepsilon_t^4} = \frac{C(t)}{2ktG^{-1}(t)} + C(t)\frac{\log t}{k(g\circ G^{-1}(t))G^{-1}(t)^2}.$$

As $G^{-1}(t)$ is a non-decreasing function and because of the hypothesis on k, the first term converges to 0 when t goes to the infinity. For the second term, we recall that $\log G(t)/g(t)$ is bounded and so,

$$G(t)^{2m(t)/k}\log G(t)/(g(t)t^2) \underset{t\to\infty}{\longrightarrow} 0,$$

because $G(t) = o(t^2)$. Lemma 3.9 asserts that if ε satisfies $\int^\infty \varepsilon_t^2 \frac{\mathrm{d}t}{C(t)} = \infty$, and $\frac{\dot{\varepsilon}(t)}{\varepsilon^5(t)} \underset{t\to\infty}{\longrightarrow} 0$, then $\lim_{t\to\infty} H(p_t|\Pi_{t,\varepsilon_t}) = 0$. For $\varepsilon_t^2 = k/\log t$ with the given condition on the constant k, we meet the required conditions and the result follows. ∎

Let us now prove Proposition 3.8 and Lemma 3.10.

Proof of Proposition 3.8. To shorten notation, let $p_t := p(t_0, x_0, t, \cdot)$ be the distribution law of the process Z_t, knowing that $Z_{t_0} = x_0$. We recall that the family of probability measures $(\Pi_{t,\varepsilon_t}, t \geq 0)$ satisfies a logarithmic Sobolev inequality $LSI(C(t))$. We also have $\Pi_{t,\varepsilon_t}(\mathrm{d}x) = \pi_{t,\varepsilon_t}(x)\lambda(\mathrm{d}x)$. Define h_t, such that $\int h_t^2 \mathrm{d}\Pi_{t,\varepsilon_t} = 1$:

$$h_t = \sqrt{\frac{p_t}{\pi_{t,\varepsilon_t}}}.$$

By Corollary 3.6, there exists a constant $C(t)$ such that

$$H(p_t|\Pi_{t,\varepsilon_t}) = \int p_t \log \frac{p_t}{\pi_{t,\varepsilon_t}} \mathrm{d}\lambda \leq C(t) \int |\nabla h_t|^2 \mathrm{d}\Pi_{t,\varepsilon_t}. \tag{3.5}$$

We now have to compute the derivative of h_t:

$$\nabla h_t = \frac{1}{2}\sqrt{\frac{p_t}{\pi_{t,\varepsilon_t}}} \left(\frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon_t^2} \right).$$

We put this last estimate in the preceding inequality (3.5) and thus

$$H(p_t|\Pi_{t,\varepsilon_t}) \leq \frac{C(t)}{4} \int p_t \left| \frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon_t^2} \right|^2 \mathrm{d}\lambda.$$

We recall that we are looking for an inequality including the time-derivative of the free energy H. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi_{t,\varepsilon_t}) = \int \dot{p}_t \log \frac{p_t}{\pi_{t,\varepsilon_t}} \mathrm{d}\lambda - \int p_t \frac{\dot{\pi}_{t,\varepsilon_t}}{\pi_{t,\varepsilon_t}} \mathrm{d}\lambda. \tag{3.6}$$

Our strategy is to find a upper bound for the two terms on the right hand side. The Kolmogorov forward equation reads

$$\dot{p}_t = \frac{1}{2}\varepsilon_t^2 \Delta p_t + Div(p_t \nabla V_t) = \nabla \cdot \left(\frac{1}{2}\varepsilon_t^2 \nabla p_t + p_t \nabla V_t \right). \tag{3.7}$$

We also remark that we have the following estimates:

$$\frac{\dot{\pi}_{t,\varepsilon_t}}{\pi_{t,\varepsilon_t}} = 4\frac{\dot{\varepsilon}(t)}{\varepsilon_t^3} \left(V_t - < V_t >_{\pi_{t,\varepsilon_t}} \right) - \frac{2}{\varepsilon_t^2} (\dot{V}_t - < \dot{V}_t >_{\Pi_{t,\varepsilon_t}}), \tag{3.8}$$

where we have used the usual notation $< f >_{\Pi_{t,\varepsilon_t}} = \int f \mathrm{d}\Pi_{t,\varepsilon_t}$. Moreover, we also find

$$\frac{\nabla \pi_{t,\varepsilon_t}}{\pi_{t,\varepsilon_t}} = -2\frac{\nabla V_t}{\varepsilon_t^2}. \tag{3.9}$$

Now put the first estimate (3.8), as well as the Kolmogorov equation (3.7), in the formula (3.6). We integrate by parts and use the logarithmic Sobolev inequality (3.5) to get

$$\begin{aligned} \int \log \frac{p_t}{\pi_{t,\varepsilon_t}} \dot{p}_t \mathrm{d}\lambda &= \int \log \frac{p_t}{\pi_{t,\varepsilon_t}} \nabla \cdot \left(\frac{1}{2}\varepsilon_t^2 \nabla p_t + p_t \nabla V_t \right) \mathrm{d}\lambda \\ &= -\int \left(\frac{\nabla p_t}{p_t} - \frac{\nabla \pi_{t,\varepsilon_t}}{\pi_{t,\varepsilon_t}} \right) \left(\frac{1}{2}\varepsilon_t^2 \nabla p_t + p_t \nabla V_t \right) \mathrm{d}\lambda \\ &= -\int \left(\frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon_t^2}, \frac{1}{2}\varepsilon_t^2 \nabla p_t + p_t \nabla V_t \right) \mathrm{d}\lambda = -\frac{\varepsilon_t^2}{2} \int p_t \left| \frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon_t^2} \right|^2 \mathrm{d}\lambda \\ &\leq -\frac{2}{C(t)} \varepsilon_t^2 H(p_t|\Pi_{t,\varepsilon_t}). \end{aligned}$$

On the other hand, we obtain the following equality for the second integral involved in the time-derivative of H:

$$\int p_t \frac{\dot{\pi}_{t,\varepsilon_t}}{\pi_{t,\varepsilon_t}} \mathrm{d}\lambda = 4\frac{\dot{\varepsilon}(t)}{\varepsilon_t^3} \int p_t(V_t - < V_t >_{\Pi_{t,\varepsilon_t}})\mathrm{d}\lambda - \frac{2}{\varepsilon_t^2} \int p_t(\dot{V}_t - < \dot{V}_t >_{\Pi_{t,\varepsilon_t}})\mathrm{d}\lambda.$$

We put all the pieces together and this leads to the result. □

Proof of Lemma 3.10. Let K be the compact set $K := \{x; V(x) \leq \eta\}$ where η is a given positive constant. As Π_{t,ε_t} converges weakly to Π_0, we only need to prove that $< |x|^2 \mathbb{1}_{K^c} >_{\Pi_{t,\varepsilon_t}}$ is bounded. We have

$$\int_{K^c} |x|^2 e^{-2\varepsilon_t^{-2} V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon_t^2}} \mathrm{d}x \leq \int_{K^c} |x|^2 e^{-2V(x)} e^{-2V(x)(\varepsilon_t^{-2}-1)} \mathrm{d}x \leq \int_{K^c} |x|^2 e^{-2V(x)} \mathrm{d}x e^{-2\eta(\varepsilon_t^{-2}-1)}.$$

By Proposition 3.2, we know that $\pi(t,\varepsilon_t) \underset{t\to+\infty}{\sim} \sum_i (2\pi\varepsilon_t^2)^{d/2} (\det \nabla^2 V(x_i))^{-1/2}$, and so there exists a positive constant $\widetilde{C}$ such that

$$< |x|^2 \mathbb{1}_{K^c} >_{\Pi_{t,\varepsilon_t}} \leq \widetilde{C} \varepsilon_t^{-d} e^{-2\eta\varepsilon_t^{-2}} \underset{t\to+\infty}{\to} 0. \quad \square$$

We will now describe the law of the limit process Y_∞.

Proposition 3.11. *The speed of convergence of $H(p(t_0, x_0, t, \cdot)|\Pi_{t,\varepsilon_t})$ toward 0 is $1/(G^{-1}(t) \log t)$.*

Proof. By Lemma 3.9, the speed of convergence is given by $\int_0^t \alpha_s e^{-\int_s^t \beta_u \mathrm{d}u} \, \mathrm{d}s$, with $\beta_s = s^{2m/k}/\log s$ and $\alpha_s = (sG^{-1}(s))^{-1} + \log s (g \circ G^{-1}(s) G^{-1}(s)^2)^{-1}$. Integrating by part, we find that $\int_0^t \beta_s \mathrm{d}s$ is equivalent, when t goes to the infinity, to $t^{1+2m/k}/\log t$ and thus, the speed of convergence is of order $(G^{-1}(t)\log t)^{-1} + (g \circ G^{-1}(t)(G^{-1}(t))^2)^{-1}$. Finally, the speed of convergence of the relative Kullback information to zero is $(G^{-1}(t)\log t)^{-1} = o((\log t)^{-1})$. □

Remark 3.2. *It is known since the work of Freidlin and Wentzell* [6]*, that the Gibbs measure Π_{t,ε_t} satisfies a large deviation principle. Therefore, the speed of convergence of Π_{t,ε_t} toward Π_0 is exponential ($e^{-\log t/2k} = t^{-1/2k}$).*

Corollary 3.12. *Suppose that $\varepsilon_t^2 = k/\log t$, where $k > 2m$. Then the process Z converges in distribution to a random variable which concentrates on the global minima of V. Thus, the process Y converges in distribution to a random variable Y_∞, which concentrates on the global minima of V.*

Proof. The Kullback information $H(p_t|\Pi_{t,\varepsilon_t})$ estimates the distance between p_t and Π_{t,ε_t}, as it is recalled in (2.5). The result follows as Π_{t,ε_t} converges weakly to Π_0. □

Remark 3.3. *The function ε is supposed to decrease slowly to zero. This is why we obtain the convergence of Y to the global minima of V. But if ε goes too fast to zero, that is $\lim_{t\to\infty} g(t)^{-1} \log G(t) = k$ with $k \leq 2m$, then Y may freeze in a local minimum. So, X does not converge in that case.*

3.2. **Study of X.** We give necessary and sufficient conditions for the convergence in distribution of X. As usual, we start to work with the process $Y_t = X_t - \overline{\mu}_t$. In order to link this section with the preceding one, we recall that $\varepsilon_t^2 = (g \circ G^{-1}(t))^{-1} = k/\log t$. It implies that we consider functions g such that (asymptotically) $\log G(t) = kg(t)$.

Let us first recall a former result.

Theorem 3.13. *(Chambeu-Kurtzmann [4, Thm5.5]) The process Y satisfies the pointwise ergodic theorem. This means that a.s., the empirical measure of Y converges weakly to a random measure, which is a convex combination of Dirac measures taken in the minimal points of V. More precisely, there exist $a_i \geq 0$ such that*

$$\frac{1}{t}\int_0^t \delta_{Y_s} \mathrm{d}s \underset{t\to\infty}{\longrightarrow} \sum_{i=1}^n a_i \delta_{m_i} \ \ a.s.$$

We are now able to conclude the study of the asymptotic behaviour of the process X.

Theorem 3.14. *Suppose that $\lim_{t\to\infty} g(t)^{-1}\log G(t) = k > 2m$. Then one of the following holds:*

1. *If V is a function such that $\sum_{1\leq i\leq n} a_i m_i = 0$, then X_t converges in distribution to $Y_\infty + \int_0^\infty Y_s \frac{\mathrm{d}s}{s}$;*
2. *Else, X_t diverges.*

Proof. Suppose that V is such that the integral $\int_0^t Y_s \frac{\mathrm{d}s}{s}$ converges a.s. The celebrated Slutsky theorem asserts that for two sequences (U_t), (W_t) of $\mathbb{R}^d$ valued random variables, if $U_t \overset{(d)}{\underset{t\to\infty}{\longrightarrow}} U$ and $|U_t - W_t| \overset{\mathbb{P}}{\underset{t\to\infty}{\longrightarrow}} 0$, then $W_t \overset{(d)}{\underset{t\to\infty}{\longrightarrow}} U$. To prove the result, we let $U_t = \overline{\mu}_t = \int_0^t \frac{\mathrm{d}s}{s} Y_s = \frac{1}{t}\int_0^t Y_s \mathrm{d}s + \int_0^t \frac{1}{s^2}\int_0^s Y_u \mathrm{d}u\mathrm{d}s$, $W_t = X_t$ and remark that

$$X_t = Y_t + \int_0^t \frac{\mathrm{d}s}{s} Y_s = Y_t + \overline{\mu}_t.$$

Suppose that V is such that $\sum_{1\leq i\leq n} a_i m_i = 0$. By Theorem 3.13, we have that $\frac{1}{t}\int_0^t Y_s \mathrm{d}s \overset{a.s.}{\longrightarrow} 0$, and we now need to find the rate of convergence in order to conclude the proof. Moreover, by [4, Prop5.3], we know that the speed of convergence of the empirical mean of the time-changed process $Y_{G^{-1}(t)}$ is $G^{-1}(1+t) - G^{-1}(t)$. But we are looking for the speed of convergence for Y_t itself. By an integration by part, we obtain that

$$\begin{aligned}\frac{1}{t}\int_0^t Y_s \mathrm{d}s &= \frac{1}{t}\int_0^{G(t)} Y_{G^{-1}(u)} \frac{\mathrm{d}u}{g\circ G^{-1}(u)} \\ &= \frac{1}{tg(t)}\int_0^{G(t)} Y_{G^{-1}(u)}\mathrm{d}u + \frac{1}{t}\int_0^{G(t)} \mathrm{d}u \frac{g'\circ G^{-1}(u)}{(g\circ G^{-1}(u))^3}\int_0^u Y_{G^{-1}(s)}\mathrm{d}s.\end{aligned}$$

Corollary 3.12 implies that the first right-hand term converges in distribution to 0 because $G(t) \leq tg(t)$. So it converges in probability to 0. It remains to prove the convergence of the second term. We recall that, up to a multiplicative positive constant, $g\circ G^{-1}(u) = \log(2+u)$. Moreover, we also know that $\frac{1}{u}\int_0^u Y_{G^{-1}(s)}\mathrm{d}s$ is a.s. bounded. So, the second right-hand term is upper bounded (up to a multiplicative positive constant) by

$$\frac{1}{t}\int_0^{G(t)} \frac{\mathrm{d}u}{(\log(2+u))^2} = \frac{G(t)}{t(\log G(t))^2} + o\left(\frac{G(t)}{t(\log G(t))^2}\right) \leq \frac{g(t)}{\log G(t)}\frac{1}{\log G(t)}$$

and the result follows: $\bar{\mu}_t = U_t$ converges in distribution. And by Corollary 3.12, Y_t converges in distribution to Y_∞ which law concentrates on the global minima of V. So, $Y_t = U_t - W_t$ converges in probability to 0.

To conclude, if V satisfies $\sum_{1\leq i\leq n} a_i m_i \neq 0$ then $\bar{\mu}_t = \int_0^t Y_s \frac{\mathrm{d}s}{s}$ does not converge and so X_t diverges. □

4. Convergence in distribution of X to a random variable

In this Section, we will prove that if g converges to 1 or 0 slowly enough, then the process X converges in distribution to an identified limit. We will first study the case $g \equiv 1$ and prove rigorously the convergence of $\bar{\mu}_t$. Then, we will consider the case $g(t) \to 0$ and $tg(t) \to +\infty$. The proof of the convergence of $\bar{\mu}_t$ will be exactly the same as in the case $g = 1$ and so, we will not reproduce it. Nevertheless, the convergence of Y will be interesting and §4.2 is essentially devoted to its proof.

4.1. If g converges toward a positive constant.

In this part, we suppose that g converges toward a positive constant, so that its primitive G goes to the infinity. In that case, we will show that the asymptotic behaviour of Y is very close to the behaviour of ξ, solution to

$$\mathrm{d}\xi_t = \mathrm{d}B_t - g(t)\nabla V(\xi_t)\mathrm{d}t.$$

Without any loss of generality, we suppose that $g(t) = 1$ for t large enough. Actually, we will prove that Y converges toward a random variable of law $\Pi(\mathrm{d}x) = \frac{e^{-2V(x)}}{\pi}\mathrm{d}x$. (Remark that the normalization constant π is well-defined as V is strictly convex out of a compact set.) To this aim, we will use the exponential decrease to zero of the relative Kullback information between the law of $Y_{G^{-1}(t)}$ and Π. Once this is done, we study the convergence of the mean $\frac{1}{t}\int_0^t Y_s \mathrm{d}s$. Indeed, we will prove that the latter integral converges if and only if $\int x \mathrm{d}\Pi(x) = 0$.

Theorem 4.1. *$X_t - X_0$ converges in distribution to Y_∞ if and only if $\int xe^{-2V(x)}\mathrm{d}x = 0$. In that case, Y_∞ has the distribution law $\frac{e^{-2V}}{\pi}$.*

The proof of this statement will be decomposed into several propositions and lemmas. We first present them all, postponing their proofs. Then, we deduce from them Theorem 4.1. Finally, we prove these intermediate results. Let us state the first of the propositions mentioned, the one showing that the time-shifted process $(Y_{G^{-1}(t)})_t$, and so $(Y_t)_t$ converges in distribution.

Proposition 4.2. *The process $(Y_t)_t$ converges in distribution to a random variable Y_∞. The distribution law of Y_∞ is $\frac{e^{-2V}}{\pi}$.*

Next, we have to show that either $\bar{\mu}_t = \int_0^t \frac{Y_s}{s}\mathrm{d}s$ converges a.s. toward $x = X_0$ for suitable functions V, or diverges.

Proposition 4.3. *$\bar{\mu}_t$ converges almost surely as $t \to \infty$ if and only if $\int xe^{-2V(x)}\mathrm{d}x = 0$.*

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Recall that $X_t = Y_t + \bar{\mu}_t$. Proposition 4.2 asserts that Y_t converges in distribution to Y_∞. Moreover, $\bar{\mu}_t$ converges almost surely to $\bar{\mu}_\infty = x$ iif $\int xe^{-2V(x)}\mathrm{d}x = 0$ by Proposition 4.3. So, we use Slutsky's theorem: $\bar{\mu}_t - \bar{\mu}_\infty$ converges a.s. to 0, and Y_t converges in distribution to Y_∞, so $Y_t + (\bar{\mu}_t - \bar{\mu}_\infty)$ goes in distribution to Y_∞. □

Let us now prove Propositions 4.2-4.3.

Proof of Proposition 4.2. We will show that the process $(Y_t)_t$ converges in distribution to Y_∞. Let p_t denote the law of Y_t. By Lemma 2.3, the probability measure $\Pi = e^{-2V}/\pi$ (where π denotes the normalization constant of Π) satisfies a logarithmic Sobolev inequality $LSI(C_{LS})$. By inequality (2.6), we know that

$$H(p_t|\Pi) \le C_{LS} \int \left| \nabla \left(\sqrt{\frac{p_t}{\Pi}} \right) \right|^2 \mathrm{d}\lambda.$$

As $\nabla \left(\sqrt{\frac{p_t}{\Pi}} \right) = \sqrt{\pi p_t} \frac{e^V}{2} \left(\frac{\nabla p_t}{p_t} + 2\nabla V \right)$, we deduce that

$$H(p_t|\Pi) \le \frac{C_{LS}}{4} \int p_t \left| \frac{\nabla p_t}{p_t} + 2\nabla V \right|^2 \mathrm{d}\lambda. \tag{4.1}$$

Moreover, by definition of the relative Kullback information, it is clear that $\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi) = \int \dot{p}_t \log \left(\frac{p_t}{\Pi} \right) \mathrm{d}\lambda$. The Kolmogorov-forward equation also reads

$$\dot{p}_t = \frac{1}{2} \Delta p_t + (\nabla p_t, \nabla V) = \nabla \cdot \left(\frac{1}{2} \nabla p_t + p_t \nabla V \right), \tag{4.2}$$

and putting this last estimate in the previous time-derivative equation of H, we have:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi) &= \int \dot{p}_t \log \left(\frac{p_t}{\Pi} \right) \mathrm{d}\lambda = \int \nabla \cdot \left(\frac{1}{2} \nabla p_t + p_t \nabla V \right) \log \left(\frac{p_t}{\Pi} \right) \mathrm{d}\lambda \\ &= -\int \left(\frac{1}{2} \nabla p_t + p_t \nabla V, \frac{\nabla p_t}{p_t} + 2\nabla V \right) \mathrm{d}\lambda \\ &= -\frac{1}{2} \int p_t \left| \frac{\nabla p_t}{p_t} + 2\nabla V \right|^2 \mathrm{d}\lambda \le \frac{-2}{C_{LS}} H(p_t|\Pi). \end{aligned}$$

So, $H(p_t|\Pi)$ converges to zero exponentially fast. This means that $\|p_t - \Pi\|_{TV}^2 \le 2H(p_t|\Pi) \to 0$, that is Y_t converges in distribution toward a random variable Y_∞. The distribution law of Y_∞ is Π and the speed of convergence is exponential. □

Proof of Proposition 4.3. Let $\bar{y}_t := \frac{1}{t} \int_0^t Y_s \mathrm{d}s$. We have to show that $\bar{y}_t$ converges almost surely to $\bar{y}_\infty = \int x \Pi(\mathrm{d}x)$. First, we decompose $\bar{\mu}_t$ in the following way

$$\bar{\mu}_t = x + \int_0^t Y_s \frac{\mathrm{d}s}{s} = x + \frac{1}{t} \int_0^t Y_s \mathrm{d}s + \int_0^t \frac{1}{s^2} \int_0^s Y_u \mathrm{d}u \mathrm{d}s.$$

We then have

$$\bar{\mu}_t = x + \bar{y}_t + \int_0^t \frac{1}{s} \bar{y}_s \mathrm{d}s. \tag{4.3}$$

Let us introduce the positive recurrent Kolmogorov process $(\xi_t, t \ge 0)$, solution to $\mathrm{d}\xi_t = \mathrm{d}B_t - \nabla V(\xi_t)\mathrm{d}t$. The invariant probability measure associated to ξ is precisely Π. As ξ is pointwise ergodic, we have for all $h \in L^1(\Pi)$:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t h(\xi_s) \mathrm{d}s = \int h \mathrm{d}\Pi \, a.s. \tag{4.4}$$

with an exponential speed of convergence (see for instance [17]).

Let us now prove the almost sure convergence of $\bar{y}_t$. We have:

$$\bar{y}_{e^{t+s}} - \bar{y}_{e^t} = \int_t^{t+s} (Y_{e^u} - \bar{y}_{e^u})\mathrm{d}u = \int_0^s (\bar{y}_\infty - \bar{y}_{e^{t+u}})\mathrm{d}u + \int_0^s (Y_{e^{t+u}} - \bar{y}_\infty)\mathrm{d}u.$$

We will now need the following technical result.

Lemma 4.4. *For all $T > 0$, $\lim_{t\to\infty} \sup_{0\le s\le T} \left|\int_0^s (Y_{e^{t+u}} - \bar{y}_\infty)\mathrm{d}u\right| = 0$ almost surely.*

Assuming the validity of this statement, the process $(\bar{y}_{e^t})_t$ is an asymptotic pseudotrajectory for the flow $\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = \bar{y}_\infty - \phi_t(x)$, $\phi_0(x) = x$. The flow induced by ϕ admits a unique limit point $\{\bar{y}_\infty\}$, which is exponentially attracted. Thus, $\bar{y}_t$ converges a.s. to $\bar{y}_\infty$ (with an exponential speed of convergence).

Let us now estimate the distance between Y_{s+e^t} and ξ_{s+e^t}, knowing that $Y_{e^t} = \xi_{e^t}$. As W is strictly convex, and $\nabla\chi$ is C_χ-Lipschitz, we obtain the following inequality:

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|Y_{s+e^t} - \xi_{s+e^t}|^2 &= -(\nabla V(Y_{s+e^t}) - \nabla V(\xi_{s+e^t}), Y_{s+e^t} - \xi_{s+e^t}) - \frac{1}{s+e^t}(Y_{s+e^t}, Y_{s+e^t} - \xi_{s+e^t}) \\ &\le -(C_W - C_\chi)|Y_{s+e^t} - \xi_{s+e^t}|^2 + \frac{1}{2(s+e^t)}|\xi_{s+e^t}|^2,\end{aligned}$$

because $-2(y, y-z) \le |z|^2$. So, we have the following bound on the square-distance:

$$|Y_{s+e^t} - \xi_{s+e^t}|^2 \le e^{-2(C_W - C_\chi)s}\int_0^s e^{(C_W-C_\chi)u}|\xi_{u+e^t}|^2\frac{\mathrm{d}u}{u+e^t}.$$

Once again, the ergodicity of ξ implies that $\frac{1}{s}\int_0^s |\xi_{u+e^t}|^2\mathrm{d}u$ converges a.s. (as $s\to\infty$) to $\int |x|^2\Pi(\mathrm{d}x)$. So there exists $C > 0$ such that

$$\begin{aligned}\int_0^s e^{(C_W-C_\chi)u}|\xi_{u+e^t}|^2\frac{\mathrm{d}u}{u+e^t} &\le e^{-t}\left(\left|\int_0^s e^{(C_W-C_\chi)u}\left(|\xi_{u+e^t}|^2 - \int |x|^2\Pi(\mathrm{d}x)\right)\mathrm{d}u\right|\right. \\ &\qquad \left. + \frac{e^{(C_W-C_\chi)s}}{C_W - C_\chi}\int |x|^2\Pi(\mathrm{d}x)\right) \\ &\le Ce^{-t}e^{(C_W-C_\chi)s}\int |x|^2\Pi(\mathrm{d}x).\end{aligned}$$

And thus $|Y_{s+e^t} - \xi_{s+e^t}|^2 \le Ce^{-t}$ a.s. So

$$\int_0^{e^t(e^s-1)} \frac{Y_{v+e^t} - \xi_{v+e^t}}{v+e^t}\mathrm{d}v = O(e^{-t}). \tag{4.5}$$

To prove Proposition 4.3, we use the decomposition (4.3). It is obvious from that decomposition that if $\bar{y}_\infty = \int x\Pi(\mathrm{d}x) \neq 0$, then $\bar{\mu}_t$ does not converge and in that case $\bar{\mu}_t \sim \bar{y}_\infty \log t$. Suppose now that $\bar{y}_\infty = 0$. As

$$\bar{y}_s = \frac{1}{s}\int_0^s Y_u\mathrm{d}u = \frac{1}{s}\int_0^s (Y_u - \xi_u)\mathrm{d}u + \frac{1}{s}\int_0^s \xi_u\mathrm{d}u,$$

and by equation (4.4), there exists a positive constant a such that $\left|\frac{1}{s}\int_0^s \xi_u\mathrm{d}u\right| \le e^{-as}$, we get that $|\bar{y}_s| = O(s^{-a})$ with $a > 0$. So, the integral $\int_0^t \frac{1}{s}\bar{y}_s\mathrm{d}s$ converges a.s., implying the convergence of the empirical mean $\bar{\mu}_t$. □

Proof of Lemma 4.4. Let $t \geq 0$. We have

$$\mathrm{d}Y_{s+e^t} = \mathrm{d}B_{s+e^t} - \left(\nabla V(Y_{s+e^t}) + \frac{Y_{s+e^t}}{s+e^t}\right)\mathrm{d}s.$$

Let us prove that the drift term $\frac{Y_{s+e^t}}{s+e^t}$ is negligible for t large enough. Let $T \geq 0$. For any $0 \leq s \leq T$, we have

$$\int_0^s (Y_{e^{t+u}} - \bar{y}_\infty)\mathrm{d}u = \int_0^{e^t(e^s-1)} \frac{Y_{v+e^t} - \xi_{v+e^t}}{v+e^t}\mathrm{d}v + \int_0^{e^t(e^s-1)} \frac{\xi_{v+e^t} - \bar{y}_\infty}{v+e^t}\mathrm{d}v, \tag{4.6}$$

where $\xi_{e^t} = Y_{e^t}$. We emphasize that ξ and Y are driven by the same Brownian motion. We have already proved in equation (4.5) that the first right-hand term of (4.6) converges (exponentially fast) to 0. Let us now study the most right-hand side of (4.6). An integration by parts leads to

$$\begin{aligned}\int_0^{e^t(e^s-1)} \frac{\xi_{v+e^t} - \bar{y}_\infty}{v+e^t}\mathrm{d}v &= \frac{e^s-1}{e^s}\left(\frac{1}{e^t(e^s-1)}\int_0^{e^t(e^s-1)} \xi_{u+e^t}\mathrm{d}u - \bar{y}_\infty\right) + \\ &+ \int_0^{e^t(e^s-1)} \frac{v}{(v+e^t)^2}\left(\frac{1}{v}\int_0^v \xi_{u+e^t}\mathrm{d}u - \bar{y}_\infty\right)\mathrm{d}v.\end{aligned} \tag{4.7}$$

The ergodicity (4.4) of ξ implies directly that $\frac{1}{e^t(e^s-1)}\int_0^{e^t(e^s-1)} \xi_{u+e^t}\mathrm{d}u - \bar{y}_\infty$ converges a.s. to 0 (as $t \to \infty$), with an exponential speed of convergence. So, there exist two positive constants a, C such that a.s.

$$\begin{aligned}\left|\int_0^{e^t(e^s-1)} \frac{v}{(v+e^t)^2}\left(\frac{1}{v}\int_0^v \xi_{u+e^t}\mathrm{d}u - \bar{y}_\infty\right)\mathrm{d}v\right| &\leq e^{-2t}\int_0^{e^t(e^s-1)} \left|\frac{1}{v}\int_0^v \xi_{u+e^t}\mathrm{d}u - \bar{y}_\infty\right|\mathrm{d}v \\ &\leq Ce^{-2t}\int_0^{e^t(e^s-1)} e^{-av}\mathrm{d}v \longrightarrow_{t\to\infty} 0.\end{aligned}$$

So, $\lim_t \sup_{0\leq s\leq T}\left|\int_0^{e^t(e^s-1)} \frac{\xi_{v+e^t}-\bar{y}_\infty}{v+e^t}\mathrm{d}v\right| = 0$ a.s. Finally, putting all the pieces together we have shown that $\bar{y}_{e^t}$ is an asymptotic pseudotrajectory for the flow generated by ϕ. ☐

4.2. **If $g(t)$ goes to zero and $tg(t)$ goes to $c \in]0, +\infty]$.** The technique we adopt here is a change of scale added to a change of measure. This is useful as soon as we wish to study the asymptotic or ergodic behaviour of a non-homogeneous process, as it usually permits to "reduce" to the homogeneous case.

4.2.1. *If $tg(t)$ goes to a positive constant.* In this part, we suppose without any loss of generality that $tg(t) = 1$. Indeed, Y is solution to the SDE

$$\mathrm{d}Y_t = \mathrm{d}B_t - \frac{1}{t}(\nabla V(Y_t) + Y_t)\mathrm{d}t. \tag{4.8}$$

Theorem 4.5. *Suppose that there exists $n \geq 1$ such that $\frac{V(x)}{|x|^{2n}}$ converges to a positive constant and $g(t) = \frac{1}{t}$. Then, X_t converges in distribution to $Y_\infty + \bar{\mu}_\infty$ if and only if $\int xe^{-2V(x)}\,\mathrm{d}x = 0$ and $n > 3/2$. In that case, Y_∞ has the distribution law $e^{-2V(x)}\,\mathrm{d}x$ (up to a positive multiplicative constant).*

Proof. The proof is similar to the one of Theorem 4.1. If V is polynomial, or $V(x) = c|x|^{2n}$ (for $c > 0$), then letting $\tilde{Y}_t := \sqrt{\frac{n-1}{n}} t^{-\frac{1}{2(n-1)}} Y_{t^{\frac{n}{n-1}}}$, we find that $\tilde{Y}$ satisfies

$$\mathrm{d}\tilde{Y}_t = \mathrm{d}B_t - \left(2cn|\tilde{Y}_t|^{2n-2}\tilde{Y}_t + \frac{2n+1}{2(n-1)t}\tilde{Y}_t\right)\mathrm{d}t.$$

Now, we can approximate the potential V by $c|x|^{2n}$ (for a well-chosen $n \geq 1$) such that the studied process $\tilde{Y}$ has the same asymptotic behaviour as

$$\mathrm{d}\xi_t = \mathrm{d}B_t - \left(\nabla V(\xi_t) + \frac{2n+1}{2(n-1)t}\xi_t\right)\mathrm{d}t$$

in the sense of asymptotic pseudotrajectory. Actually, one easily shows that if $\int xe^{-2V(x)}\,\mathrm{d}x = 0$, then for all $T > 0$

$$\sup_{0\leq s\leq T} |\tilde{Y}_{s+e^t} - \xi_{s+e^t}|^2 = O(e^{-t}). \tag{4.9}$$

This has been already shown for Lemma 4.4 and we do not reproduce the proof here. Finally, this proves that $\tilde{Y}$ has the same ergodic behaviour as ξ (see [3]). So, as ξ is ergodic and almost surely $\frac{1}{t}\int_0^t \xi_s\,\mathrm{d}s$ converges, the limit points of $\frac{1}{t}\int_0^t \delta_{\tilde{Y}_s}\,\mathrm{d}s$ are included into the set of the invariant measures of ξ, that is $e^{-2V(x)}\mathrm{d}x$. So $\frac{1}{t}\int_0^t \tilde{Y}_s\,\mathrm{d}s$ goes to zero iif $\int xe^{-2V(x)}\,\mathrm{d}x = 0$. By (4.9), the asymptotic pseudotrajectory has a polynomial speed. Moreover, ξ converges to its invariant probability measure with an exponential speed. Thus, $\frac{1}{t}\int_0^t \tilde{Y}_s\,\mathrm{d}s$ converges a.s. to zero with a polynomial speed of convergence (of the order of $1/t$) if $\int xe^{-2V(x)}\,\mathrm{d}x = 0$ (and diverges otherwise). Now, remembering that $\tilde{Y}_s = s^{-\frac{1}{2(n-1)}} Y_{s^{\frac{n}{n-1}}}$, we conclude that

$$\frac{1}{t}\int_0^t \tilde{Y}_s\,\mathrm{d}s = \frac{1}{t}\int_0^{t^{\frac{n}{n-1}}} Y_u\,\frac{\mathrm{d}u}{u^{\frac{3}{2n}}} \longrightarrow 0 \text{ } a.s. \tag{4.10}$$

Indeed, we have

$$\frac{1}{t}\int_0^{t^{\frac{n}{n-1}}} Y_u\,\frac{\mathrm{d}u}{u^{\frac{3}{2n}}} = \frac{1}{t^{1+\frac{3}{2(n-1)}}}\int_0^{t^{\frac{n}{n-1}}} Y_u\,\mathrm{d}u - \frac{1}{t}\int_0^{t^{\frac{n}{n-1}}} u^{-1-\frac{3}{2n}}\int_0^u Y_s\,\mathrm{d}s\,\mathrm{d}u = O(t^{-1}) \text{ } a.s.$$

And thus, $\frac{1}{T}\int_0^T Y_u\,\mathrm{d}u$ converges a.s. to 0 (with $T = t^{\frac{n}{n-1}}$) iif $n > 3/2$. We then refer to §4.1 to obtain the convergence of the process X. □

We also remark that this result is coherent with the basic Ornstein-Uhlenbeck case.

4.2.2. *If $tg(t)$ goes to the infinity.* This study will be divided into two different cases. First, we suppose that there exists $0 < \alpha < 1$ such that $t^\alpha g(t)$ goes to a positive constant. Whereas in the second case, $t^\alpha g(t)$ goes to the infinity for any $0 < \alpha < 1$ (this is for instance satisfied by $g(t) = 1/\log t$). The first part of the study is identical for the two cases and we will only divide the end of the study.

Theorem 4.6. *Suppose that there exists $n \geq 1$ such that $\frac{V(x)}{|x|^{2n}}$ converges to a positive constant.*

1. *If there exists $0 < \alpha < 1$ such that $g(t) = t^\alpha$, then X_t converges in distribution to $Y_\infty + \bar{\mu}_\infty$ if and only if $\int xe^{-2V(x)}\,\mathrm{d}x = 0$ and $n > 4\alpha$.*
2. *If $t^\alpha g(t) \to +\infty$ for all $0 < \alpha < 1$, then X_t converges in distribution to $Y_\infty + \bar{\mu}_\infty$ if and only if $\int xe^{-2V(x)}\,\mathrm{d}x = 0$.*

Let us sketch the proof of Theorem 4.6, that is postponed to the end of the paragraph. As in the preceding paragraph, we first suppose that $V(x) = c|x|^{2n}$. Define f as the positive increasing solution to $g \circ f(t) = (f'(t))^{-n}$. Consider the time and scale-changed process $\tilde{Y}$ defined by $\tilde{Y}_t := \frac{Y_{f(t)}}{\sqrt{f'(t)}}$. Applying Itô's formula to $\tilde{Y}$, we thus find that $\tilde{Y}$ satisfies the SDE

$$\mathrm{d}\tilde{Y}_t = \mathrm{d}B_t - \left(2nc\tilde{Y}_t|\tilde{Y}_t|^{2n-2} + \left(\frac{f'(t)}{f(t)} + \frac{f''(t)}{2f'(t)}\right)\tilde{Y}_t\right)\mathrm{d}t. \tag{4.11}$$

Now, we approximate the potential V by $|x|^{2n}$ for a well-chosen $n \geq 1$. So, the studied process $\tilde{Y}$ has the same asymptotic behaviour as

$$\mathrm{d}\xi_t = \mathrm{d}B_t - \left(\nabla V(\xi_t) + \frac{\xi_t}{\beta(t)}\right)\mathrm{d}t, \tag{4.12}$$

where $\beta(t)$ is defined by $\frac{1}{\beta(t)} = \frac{f'(t)}{f(t)} + \frac{f''(t)}{2f'(t)}$ (this last quantity goes to 0 as t tends to $+\infty$). Define also the process $\hat{Y}$ as the solution to the SDE

$$\mathrm{d}\hat{Y}_t = \mathrm{d}B_t - \nabla V(\hat{Y}_t)\mathrm{d}t. \tag{4.13}$$

Lemma 4.7. *The process ξ (and also $\tilde{Y}$) is an asymptotic pseudotrajectory for the process $\hat{Y}$: for all $t, T > 0$, we have $\sup_{0\leq s\leq T} |\hat{Y}_{s+t} - \xi_{s+t}| = O(\beta(t)^{-1/2})$ a.s.*

Proof. Let $t > 0$ and $0 \leq s \leq T$. Itô's formula implies that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\hat{Y}_{s+t} - \xi_{s+t}|^2 &= -(\nabla V(\hat{Y}_{s+t}) - \nabla V(\xi_{s+t}), \hat{Y}_{s+t} - \xi_{s+t}) + \frac{1}{\beta(s+t)}(\xi_{s+t}, \hat{Y}_{s+t} - \xi_{s+t}) \\ &\leq -(C_W - C_\chi)|\hat{Y}_{s+t} - \xi_{s+t}|^2 + \frac{1}{2\beta(s+t)}|\hat{Y}_{s+t}|^2.\end{aligned}$$

So, we find that

$$|\hat{Y}_{s+t} - \xi_{s+t}|^2 \leq e^{-2(C_W-C_\chi)s}\int_0^s e^{(C_W-C_\chi)u}|\hat{Y}_{u+t}|^2\frac{\mathrm{d}u}{\beta(u+t)}.$$

Let us note $\pi = \int e^{-2V(x)}\mathrm{d}x$ the normalisation constant. The ergodicity of $\hat{Y}$ implies the existence of $C > 0$ such that

$$\begin{aligned}\int_0^s e^{(C_W-C_\chi)u}|\hat{Y}_{u+t}|^2\frac{\mathrm{d}u}{\beta(u+t)} &\leq \frac{1}{\beta(t)}\left\{\left|\int_0^s e^{(C_W-C_\chi)u}\left(|\hat{Y}_{u+t}|^2 - \int |x|^2\frac{e^{-2V(x)}}{\pi}\,\mathrm{d}x\right)\mathrm{d}u\right|\right. \\ &\quad \left. + \frac{e^{(C_W-C_\chi)s}}{C_W - C_\chi}\int |x|^2\frac{e^{-2V(x)}}{\pi}\,\mathrm{d}x\right\} \\ &\leq \frac{C}{\beta(t)}e^{(C_W-C_\chi)s}\int |x|^2 e^{-2V(x)}\,\mathrm{d}x.\end{aligned}$$

This leads to the result. And also there exists a constant $M > 0$ such that a.s.

$$\sup_{0\leq s\leq T}|\hat{Y}_{s+t} - \xi_{s+t}|^2 \leq \frac{M}{\beta(t)}. \tag{4.14}$$

□

Proof of Theorem 4.6. Lemma 4.7 proves that $\frac{1}{t}\int_0^t \tilde{Y}_u\,\mathrm{d}u$ converges a.s. to 0 iif $\int xe^{-2V(x)}\,\mathrm{d}x = 0$. Moreover, we see by a similar Eq. (4.14) that $\tilde{Y}$ is an asymptotic pseudotrajectory for $\hat{Y}$ with the speed of convergence $\frac{1}{\sqrt{\beta(t)}}$. As $\hat{Y}$ converges to its invariant probability measure

$e^{-2V(x)}\mathrm{d}x$ with an exponential speed of convergence, we find that $\frac{1}{t}\int_0^t \tilde{Y}_s \,\mathrm{d}s$ converges a.s. to 0 iif $\int xe^{-2V(x)}\,\mathrm{d}x = 0$ and in that case, we have the following result, depending on the function g.

1) First, suppose that $g(t) = t^{-\alpha}$ for a given $0 < \alpha < 1$. We thus have $f(t) = t^{\frac{n}{n-\alpha}}$ and $\frac{1}{t}\int_0^t Y_{s^{\frac{n}{n-\alpha}}} \frac{\mathrm{d}s}{s^{\frac{\alpha}{2(n-\alpha)}}} = \frac{1}{t}\int_0^{t^{\frac{n}{n-\alpha}}} Y_u \frac{\mathrm{d}u}{u^{\frac{3\alpha}{2n}}}$ converges to 0 iif $\int xe^{-2V(x)}\,\mathrm{d}x = 0$. Then, a.s. $\left|\frac{1}{T}\int_0^T Y_u \,\mathrm{d}u\right|$ behaves asymptotically as $T^{\frac{2\alpha}{n}-\frac{1}{2}}$, where $T = t^{\frac{n}{n-\alpha}}$ and thus it converges to zero iif $n > 4\alpha$.

2) Suppose now that $t^\alpha g(t) \to \infty$ for all $0 < \alpha < 1$. This implies that $g(t) \geq t^{-\alpha}$ for any $0 < \alpha < 1$ and so $f(t) \leq t^{\frac{n}{n-\alpha}}$ and $0 \leq \frac{-g'(t)}{g(t)} \leq \frac{\alpha}{t}$. So

$$\frac{1}{t}\int_0^t \tilde{Y}_s \mathrm{d}s = \frac{1}{t}\int_{f(0)}^{f(t)} Y_u \frac{\mathrm{d}u}{(f' \circ f^{-1}(u))^{3/2}}$$

and for $T = f(t)$, the mean $\left|\frac{1}{T}\int_0^T Y_u \,\mathrm{d}u\right|$ is a.s. upper bounded by $\frac{f(t)}{t(f'(t))^{3/2}\sqrt{\beta(t)}}$. As

$$\frac{1}{\beta(t)} = \frac{f'(t)}{f(t)} + \frac{f''(t)}{2f(t)} \leq \frac{(g \circ f(t))^{-1/n}}{f} + \frac{\alpha}{nf(t)(g \circ f(t))^{2/n}} \leq \frac{1}{\sqrt{f(t)}(g \circ f(t))^{1/n}},$$

we find that

$$\frac{f(t)}{t(f'(t))^{3/2}\sqrt{\beta(t)}} \leq \frac{\sqrt{f(t)}(g \circ f(t))^{\frac{1}{2n}}}{t} \leq (g \circ f(t))^{\frac{1}{2n}} t^{\frac{n}{2(n-\alpha)}-1}.$$

This last term goes to 0 if $n > 2\alpha$. As α is arbitrary chosen between 0 and 1, we conclude that $\left|\frac{1}{T}\int_0^T Y_u \,\mathrm{d}u\right|$ converges a.s. to 0 for any $n \geq 1$. □

Acknowledgement: I wish to thank an anonymous referee for careful reading and helpful comments.

References

[1] Bakry D. & Emery M. [1985], Diffusions hypercontractives, *Sém. Prob. XIX*, Springer LNM 1123, 177-206.

[2] Benaïm M. & Hirsch M.W. [1996], Asymptotic pseudotrajectories and chain reccurent flows, *J. Dynam. Diff. Equ.* **8**, 141-176.

[3] Benaïm M. & Schreiber S.J. [2000], Weak asymptotic pseudotrajectories for semiflows: ergodic properties, *J. Dynamics and Diff. Eq.* **12**(3), 579-598.

[4] Chambeu S. & Kurtzmann A. [2011], Some particular self-interacting diffusions: ergodic behaviour and almost sure convergence, *Bernoulli*, **17**(4), 1248-1267.

[5] Chiang T.S., Hwang C.R. & Sheu S.J. [1987], Diffusion for global optimization in $\mathbb{R}^n$, *SIAM J. Control Optim.*, **25**, 737-753.

[6] Freidlin M.I. & Wentzell A.D. [1984], *Random perturbation of dynamical systems*, Springer.

[7] Gradinaru M. & Offret Y. [2012], Existence and asymptotic behaviour of some time-inhomogeneous diffusions, *Ann. I.H.P Prob. Stat.*, to appear.

[8] Hajek B. [1988], Cooling schedules for optimal annealing, *Math. Oper. Res.*, **13**, 311-329.

[9] Holley R.A., Kusuoka S. & Stroock D.W. [1989], Asymptotics of the spectral gap with applications to the theory of simulated annealing, *J. Funct. Anal.*, **83**, 333-347.

[10] Holley R.A. & Stroock D.W. [1987], Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Stat. Phys.* **46**, 1159-1194.

[11] Holley R.A. & Stroock D.W. [1988], Simulated annealing via Sobolev inequalities, *Comm. Math. Phys.*, 553-569.

[12] Hwang C.R. [1980], Laplace's method revisited: weak convergence of probability measures, *Ann. Proba.*, **8**, 1177-1182.

[13] Ivanov A.F., Kazmerchuk Y.I. & Swishchuk A.V. [2003], Theory, stochastic stability and applications of stochastic delay differential equations: a survey of results, *Differential Equations Dynam. Systems*, **11** (1-2), 55-115.

[14] Jacot S. [1992], Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov, *J. Mult. An.*, **40**, 335-347.

[15] Miclo L. [1992], Recuit simulé sur $\mathbb{R}^n$. Étude de l'évolution de l'énergie libre, *Ann. Inst. H. Poincaré*, **28** (2), 235-266.

[16] Reed M. & Simon B. [1978], *Methods of Modern Mathematical Physic*, Vol.IV, Academic Press, New-York London.

[17] Revuz D. & Yor M. [1998], *Continuous Martingales and Brownian Motion*, 3rd edition, Springer.

Aline Kurtzmann: Université de Lorraine, Institut Elie Cartan Lorraine, UMR 7502 CNRS, Vandœuvre-lès-Nancy, F-54506, France. *aline.kurtzmann@univ-lorraine.fr*