Non-asymptotic convergence analysis for the Unadjusted Langevin Algorithm

Alain Durmus, Éric Moulines

Non-asymptotic convergence analysis for the Unadjusted Langevin Algorithm

Alain Durmus [*] **and Éric Moulines** [**]

LTCI, Telecom ParisTech & CNRS,
46 rue Barrault, 75634 Paris Cedex 13, France.
e-mail: [*]alain.durmus@telecom-paristech.fr

Centre de Mathématiques Appliquées, UMR 7641,
Ecole Polytechnique,
route de Saclay, 91128 Palaiseau cedex, France.
e-mail: [**]eric.moulines@polytechnique.edu

Abstract: Sampling distribution over high-dimensional state-space is a problem which has recently attracted a lot of research efforts; applications include Bayesian non-parametrics, Bayesian inverse problems and aggregation of estimators. All these problems boil down to sample a target distribution π having a density w.r.t. the Lebesgue measure on $\mathbb{R}^d$, known up to a normalisation factor (see [9] for details) $x \mapsto \mathrm{e}^{-U(x)}/\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\mathrm{d}y$ where U is continuously differentiable with Lipschitz gradient. In this paper, we study a sampling technique based on the Euler discretization of the Langevin stochastic differential equation. Contrary to the Metropolis Adjusted Langevin Algorithm (MALA), we do not apply a Metropolis-Hastings correction. We obtain for both constant and decreasing step sizes in the Euler discretization, non-asymptotic bounds for the convergence to the target distribution π in total variation distance. A particular attention is paid on the dependence on the dimension of the state space, to demonstrate the applicability of this method in the high dimensional setting, at least when U is convex. These bounds improve and extend the results of [10].

AMS 2000 subject classifications: primary 65C05, 60F05, 62L10; secondary 65C40, 60J05,93E35.
Keywords and phrases: total variation distance, Langevin diffusion, Markov Chain Monte Carlo, Metropolis Adjusted Langevin Algorithm, Rate of convergence.

1. Introduction

We study the sampling over $\mathbb{R}^d$ of a target distribution π with density $x \mapsto \mathrm{e}^{-U(x)}/\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\mathrm{d}y$ w.r.t. the Lebesgue measure, where U is a continuously differentiable function. We consider a sampling method based on the Euler discretization of the Langevin stochastic differential equation (SDE)

$$\mathrm{d}Y_t^{\mathrm{L}} = -\nabla U(Y_t^{\mathrm{L}})\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^d\,, \tag{1}$$

where $(B_t^d)_{t\geq 0}$ is a d-dimensional Brownian motion. It is well-known that the Markov semi-group associated with the Langevin diffusion $(Y_t^{\mathrm{L}})_{t\geq 0}$ is reversible w.r.t. π. Under suitable conditions, the convergence to π takes place at geometrical rate. Precise quantitative estimates of the rate of convergence with explicit dependence on the dimension d of the state space have been recently obtained using either functional inequalities (such as Poincaré and log-Sobolev inequalities; see [2, 7] [3]) or by coupling techniques (see [12]). The Euler-Maruyama discretization scheme associated to the Langevin diffusion defines the discrete time-Markov chain given by

$$X_{k+1} = X_k - \gamma_{k+1}\nabla U(X_k) + \sqrt{2\gamma_{k+1}}Z_{k+1} \tag{2}$$

where $(Z_k)_{k\geq 1}$ is an i.i.d. sequence of standard Gaussian random variables and $(\gamma_k)_{k\geq 1}$ is a sequence of step sizes, which can either be held constant or be chosen to decrease to 0. The

idea of using the Markov chain $(X_n)_{n\geq 0}$ to sample (approximately) from the target π has been first introduced in the physics literature by [30] and popularized in the computational statistics community by [14] and [15]. It has been studied in depth by [31], which proposed to use a Metropolis-Hastings step at each iteration to enforce reversibility w.r.t. π leading to the Metropolis Adjusted Langevin Algorithm (MALA). They coin the term *unadjusted* Langevin algorithm (ULA) when the Metropolis-Hastings step is avoided.

The purpose of this paper is to study the convergence of the ULA algorithm. The emphasis is put on non-asymptotic computable bounds; we pay a particular attention to the way these bounds scale with the dimension d and constants characterizing the smoothness and curvature at infinity of the potential U. Our study covers both the constant and decreasing step sizes and we analyse both the "finite horizon" (where the total number of simulations is specified before running the algorithm) and "any-time" settings (where the algorithm can be stopped after any iteration).

When the step size $\gamma_k = \gamma$ is constant, under appropriate conditions related to [31], the Markov chain $(X_n)_{n\geq 0}$ is V-uniformly geometrically ergodic with stationary distribution π_γ. With few exceptions, the stationary distribution π_γ is no longer equal to the target π. It is of course sensible to expect that if the step size γ is small enough, then the stationary distribution of this chain is in some sense close to π. We provide non-asymptotic bounds of the V-total variation distance between π_γ and π, with explicit dependence on the step size γ, the dimension d and the number of simulations. Our results complete and extend the recent works by [11] and [10]. We first consider the case of densities which are super-exponential in the tails, for which we prove convergence in total variation. We also address the case where π is globally log-concave and strongly log-concave outside a ball.

When $(\gamma_k)_{k\geq 1}$ decreases to zero, then $(X_n)_{n\geq 0}$ is a non-homogeneous Markov chain. If in addition $\sum_{k=1}^{\infty} \gamma_k = \infty$, we show that the marginal distribution of this non-homogeneous chain converges, under some mild additional conditions, to the target distribution π. We also provide explicit expression for the convergence rate, emphasizing the role of the dimension and the rate of decrease of the sequence $(\gamma_k)_{k\geq 1}$. Compared to the related works by [20], [21], [22] and [23], we establish not only the weak convergence of the weighted empirical measure of the path to the target distribution, but establish a much stronger convergence in total variation, similarly to [10], where the strongly log-concave case is considered.

The paper is organized as follows. In Section 2, the main convergence results are stated under abstract assumptions. We then specialize these results to different classes of densities; in Section 2.1, these results are stated for a class of densities which are superexponential in the tails. In Section 2.2 these results are sharpened by considering densities which are log-concave. Finally, in Section 2.3, densities which are log-concave and strongly log-concave in the tails are dealt with. The proofs are gathered in Section 3. Some general convergence results for diffusion based on reflection coupling, which are of independent interest, are stated in Section 4.

Notations and conventions

Denote by $\mathcal{B}(\mathbb{R}^d)$ the Borel σ-field of $\mathbb{R}^d$ and by $\mathbb{F}(\mathbb{R}^d)$ the set of all Borel measurable functions on $\mathbb{R}^d$ and for $f \in \mathbb{F}(\mathbb{R}^d)$, $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|$. Denote by $\mathbb{M}(\mathbb{R}^d)$ the space of finite signed measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $\mathbb{M}_0(\mathbb{R}^d) = \{\mu \in \mathbb{M}(\mathbb{R}^d) \mid \mu(\mathbb{R}^d) = 0\}$. For $\mu \in \mathbb{M}(\mathbb{R}^d)$ and $f \in \mathbb{F}(\mathbb{R}^d)$ a μ-integrable function, denote by $\mu(f)$ the integral of f w.r.t. μ. Let $V : \mathbb{R}^d \to [1, \infty)$ be a measurable function. For $f \in \mathbb{F}(\mathbb{R}^d)$, the V-norm of f is given by $\|f\|_V = \sup_{x\in\mathbb{R}^d} |f(x)|/V(x)$.

For $\mu \in \mathbb{M}(\mathbb{R}^d)$, the V-total variation distance of μ is defined as

$$\|\mu\|_V = \sup_{f \in \mathbb{F}(\mathbb{R}^d), \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f(x) \mathrm{d}\mu(x) \right|$$

If $V \equiv 1$, then $\|\cdot\|_V$ is the total variation denoted by $\|\cdot\|_{\mathrm{TV}}$.

For $p \geq 1$, denote by $\mathrm{L}^p(\pi)$ the set of measurable functions such that $\pi(|f|^p) < \infty$. For $f \in \mathrm{L}^2(\pi)$, the variance of f under π is denoted by $\mathrm{Var}_\pi\{f\}$. For all function f such that $f \log(f) \in \mathrm{L}^1(\pi)$, the entropy of f with respect to π is defined by

$$\mathrm{Ent}_\pi(f) = \int_{\mathbb{R}^d} f(x) \log(f(x)) \mathrm{d}\pi(x) \ .$$

Let μ and ν be two probability measures on $\mathbb{R}^d$. If $\mu \ll \nu$, we denote by $\mathrm{d}\mu/\mathrm{d}\nu$ the Radon-Nikodym derivative of μ w.r.t. ν. For $k \geq 0$, denote by $C^k(\mathbb{R}^d)$, the set of k-times continuously differentiable functions. For $f \in C^2(\mathbb{R}^d)$, denote by ∇f the gradient of f and Δf the Laplacian of f. For all $x \in \mathbb{R}^d$ and $M > 0$, we denote by $\mathrm{B}(x, M)$, the ball centered at x of radius M. In the sequel, we take the convention that for $n, p \in \mathbb{N}$, $n < p$ then $\sum_p^n = 0$ and $\prod_p^n = 1$.

2. Main results

In this Section, we will present our main results. We also outline the main ideas underlying the proofs. Consider the following assumption on the potential U:

L 1. *The function U is continuously differentiable on $\mathbb{R}^d$ and is gradient Lipschitz, i.e. there exists $L \geq 0$ such that for all $x, y \in \mathbb{R}^d$,*

$$\|\nabla U(x) - \nabla U(y)\| \leq L \|x - y\| \ .$$

Under **L**1, by [17, Theorem 2.4-3.1] for every initial point $x \in \mathbb{R}^d$, there exists a unique strong solution $(Y_t^{\mathrm{L}}(x))_{t \geq 0}$ to the Langevin SDE (1). Define for all $t \geq 0$, $x \in \mathbb{R}^d$ and $\mathrm{A} \in \mathcal{B}(\mathbb{R}^d)$, $P_t^{\mathrm{L}}(x, \mathrm{A}) = \mathbb{P}\left(Y_t^{\mathrm{L}}(x) \in \mathrm{A}\right)$. The semi-group $(P_t^{\mathrm{L}})_{t \geq 0}$ is reversible w.r.t. π, and hence admits π as its (unique) invariant measure. In the sequel, we consider the case where $(P_t^{\mathrm{L}})_{t \geq 0}$ is uniformly geometrically ergodic, *i.e.* there exists $\kappa \in [0, 1)$ such that for any initial distribution μ_0 and $t > 0$,

$$\left\| \mu_0 P_t^{\mathrm{L}} - \pi \right\|_{\mathrm{TV}} \leq C(\mu_0) \kappa^t \ , \tag{3}$$

for some computable constant $C(\mu_0) \in [0, +\infty]$. Denote by $\mathscr{A}^{\mathrm{L}}$ the generator associated with the Langevin diffusion (1), given for all $f \in C^2(\mathbb{R}^d)$ by

$$\mathscr{A}^{\mathrm{L}} f = -\langle \nabla U, \nabla f \rangle + \Delta f \ . \tag{4}$$

A twice continuously differentiable function $V : \mathbb{R}^d \to [1, \infty)$ is a *Lyapunov function* for the generator $\mathscr{A}^{\mathrm{L}}$ if there exists $\theta > 0$, $\beta \geq 0$ and a measurable set $\mathcal{E}$ such that,

$$\mathscr{A}^{\mathrm{L}} V \leq -\theta V + \beta 1_{\mathcal{E}} \ . \tag{5}$$

By [31, Theorem 2.2], if V is a Lyapunov function for $\mathscr{A}^{\mathrm{L}}$, the Langevin diffusion is uniformly geometrically ergodic, with constants $C(\mu_0)$ and the rate κ which can be made explicit, but are in general pessimistic and depends exponentially on the dimension. We will see in the sequel other methods to obtain (3) to obtain tighter control on $C(\mu_0)$ and κ. These upper bounds depend crucially on the assumptions that we are ready to make on the potential U and also on the

methods of proofs (functional inequalities and coupling techniques yield different values for this constant). The dependence in the dimension d is of primary interest in all the derivations that follow.

Consider now the Euler discretization of the diffusion (2). Let $(\gamma_k)_{k\geq 1}$ be a sequence of positive and non-increasing step sizes and for $n, p \in \mathbb{N}$, denote by

$$\Gamma_{n,p} \stackrel{\text{def}}{=} \sum_{k=n}^{p} \gamma_k \;, \qquad \Gamma_n = \Gamma_{1,n} \;. \tag{6}$$

For $\gamma > 0$, consider the Markov kernel R_γ given for all $\mathrm{A} \in \mathcal{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ by

$$R_\gamma(x, \mathrm{A}) = \int_{\mathrm{A}} (4\pi\gamma)^{-d/2} \exp\left(-(4\gamma)^{-1} \left\|y - x + \gamma \nabla U(x)\right\|^2\right) \mathrm{d}y \;. \tag{7}$$

The discretized Langevin diffusion $(X_n)_{n\geq 0}$ given in (2) is a time-inhomogeneous Markov chain, for $p \geq n \geq 1$ and $f \in \mathbb{F}_+(\mathbb{R}^d)$, $\mathbb{E}^{\mathcal{F}_n}\left[f(X_p)\right] = Q_\gamma^{n,p} f(X_n)$ where $\mathcal{F}_n = \sigma(X_\ell, 0 \leq \ell \leq n)$ and

$$Q_\gamma^{n,p} = R_{\gamma_n} \cdots R_{\gamma_p} \;, \qquad Q_\gamma^n = Q_\gamma^{1,n} \;, \tag{8}$$

with the convention that for $n, p \geq 0$, $n < p$, $Q_\gamma^{p,n}$ is the identity operator. Under **L** 1, the Markov kernel R_γ is strongly Feller, irreducible, strongly aperiodic. We will say that a function $V : \mathbb{R}^d \to [1, \infty]$ satisfies a Foster-Lyapunov drift condition for R_γ if there exist constants $\bar{\gamma} > 0$, $\lambda \in [0, 1)$ and $c > 0$ such that, for all $\gamma \in (0, \bar{\gamma}]$, we get

$$R_\gamma V \leq \lambda^\gamma V + \gamma c \;. \tag{9}$$

Note the specific form of the Foster-Lyapunov condition as a function of $\gamma > 0$. It reflects how the mixing rate of the Markov chain depends upon the step size $\gamma > 0$. If $\gamma = 0$, then $R_0(x, A) = \delta_x(A)$ for $x \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$, the Markov chain is not mixing. Intuitively, as γ get larger, then the step sizes are larger and the mixing of R_γ increases. If R_γ satisfies a Foster-Lyapunov condition for some $\gamma > 0$, then R_γ admits a unique stationary distribution π_γ and that the Markov kernel is V-uniformly geometrically ergodic, *i.e.* there exist constants $C(\gamma) < \infty$ and $\lambda \in [0, 1)$, such that for all $x \in \mathbb{R}^d$,

$$\left\|R_\gamma^k(x, \cdot) - \pi_\gamma\right\|_V \leq C(\gamma) \lambda^{\gamma\, k} V(x) \;. \tag{10}$$

The constants $C(\gamma)$ and λ depend once again on the assumptions on the potential U. For reason that will become obvious in the sequel, we will not use estimates of the form (10) but rather use the Foster-Lyapunov drift to control quantitatively the moments of the time-inhomogeneous chain. The types of bounds which are needed, are summarised in the following elementary Lemma.

Lemma 1. *Let $\bar{\gamma} > 0$. Assume that for all $x \in \mathbb{R}^d$ and $\gamma \in [0, \bar{\gamma}]$, (9) holds for some constants $\lambda \in (0, 1)$ and $c > 0$. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of nonincreasing step sizes such that $\gamma_k \in [0, \bar{\gamma}]$ for all $k \in \mathbb{N}^*$. Then for all $n \geq 0$, $Q_\gamma^n V(x) \leq F(\lambda, \Gamma_{1,n}, c, \gamma_1, V(x))$ where*

$$F(\lambda, a, c, \gamma, w) = \lambda^a w + c(-\lambda^\gamma \log(\lambda))^{-1} \;. \tag{11}$$

Proof. The proof is postponed to Section 3.1. □

Note that Lemma 1 implies that $\sup_{k\geq 0}\{Q_\gamma^k V(x)\} \leq G(\lambda, c, \gamma_1, V(x))$ where

$$G(\lambda, c, \gamma, w) = w + c(-\lambda^\gamma \log(\lambda))^{-1} \;. \tag{12}$$

We give below the main ingredients which are needed to obtain a quantitative bound for $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}}$ for all $x \in \mathbb{R}^d$. This quantity is decomposed as follows: for all $0 \le n < p$,

$$\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \le \|\delta_x Q_\gamma^n Q_\gamma^{n+1,p} - \delta_x Q_\gamma^n P^{\mathrm{L}}_{\Gamma_{n+1},p}\|_{\mathrm{TV}} + \|\delta_x Q_\gamma^n P^{\mathrm{L}}_{\Gamma_{n+1},p} - \pi\|_{\mathrm{TV}} \; . \tag{13}$$

To control the first term on the right hand side, we use a method borrowed from [11]. For $0 \le s \le t$, let $\mathrm{C}([s,t],\mathbb{R}^d)$ be the space of continuous function on $[s,t]$ taking values in $\mathbb{R}^d$. For all $y \in \mathbb{R}^d$, denote by $\mu^y_{n,p}$ and $\bar\mu^y_{n,p}$ the laws on $\mathrm{C}([\Gamma_n,\Gamma_p],\mathbb{R}^d)$ of the Langevin diffusion $(Y^{\mathrm{L}}_t(y))_{\Gamma_n \le t \le \Gamma_p}$ and of the continuously-interpolated Euler discretization $(\bar Y^{\mathrm{L}}_t(y))_{\Gamma_n \le t \le \Gamma_p}$ both started at y at time Γ_n. Denote by $(Y_t(y), \overline{Y}_t(y))_{t\ge 0}$ the unique strong solution started at (y,y) at time $t = \Gamma_n$ of the time-inhomogeneous diffusion defined for $t \ge 0$ by

$$\begin{cases} \mathrm{d}Y^{\mathrm{L}}_t = -\nabla U(Y^{\mathrm{L}}_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t \\ \mathrm{d}\bar Y^{\mathrm{L}}_t = -\overline{\nabla U}(\bar Y^{\mathrm{L}}, t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t \; , \end{cases} \tag{14}$$

where $(\Gamma_n)_{n\ge 1}$ is given in (6) and for any continuous function $\mathrm{w} : \mathbb{R}_+ \to \mathbb{R}^d$, $s \mapsto \mathrm{w}_s$, and $t \ge 0$

$$\overline{\nabla U}(\mathrm{w}, t) = \sum_{k=0}^{\infty} \nabla U(\mathrm{w}_{\Gamma_k}) 1_{[\Gamma_k,\Gamma_{k+1})}(t) \; . \tag{15}$$

Girsanov's Theorem [18, Theorem 5.1, Corollary 5.16, Chapter 3] shows that $\mu^y_{n,p}$ and $\bar\mu^y_{n,p}$ are mutually absolutely continuous and in addition, $\bar\mu^y_{n,p}$-almost surely

$$\frac{\mathrm{d}\mu^y_{n,p}}{\mathrm{d}\bar\mu^y_{n,p}} = \exp\left(\frac{1}{2}\int_{\Gamma_n}^{\Gamma_p} \left\langle \nabla U(\bar Y^{\mathrm{L}}_s(y)) - \overline{\nabla U}(\bar Y^{\mathrm{L}}(y), s), \mathrm{d}\bar Y^{\mathrm{L}}_s(y) \right\rangle \right.$$
$$\left. - \frac{1}{4}\int_{\Gamma_n}^{\Gamma_p} \left\{ \left\| \nabla U(\bar Y^{\mathrm{L}}_s(y)) \right\|^2 - \left\| \overline{\nabla U}(\bar Y^{\mathrm{L}}(y), s) \right\|^2 \right\} \mathrm{d}s \right) \; . \tag{16}$$

Under **L**1, for all $y \in \mathbb{R}^d$ it yields

$$\mathrm{KL}(\mu^y_{n,p} | \bar\mu^y_{n,p}) \le 4^{-1} \int_{\Gamma_n}^{\Gamma_p} \mathbb{E}\left[\left\| \nabla U(\bar Y^{\mathrm{L}}_s(y)) - \overline{\nabla U}(\bar Y^{\mathrm{L}}(y), s) \right\|^2 \right] \mathrm{d}s$$
$$\le 4^{-1} L \sum_{k=n}^{p-1} \left\{ (\gamma_{k+1}^3/3) \int_{\mathbb{R}^d} \|\nabla U(z)\|^2 \, Q_\gamma^{n+1,k}(y, \mathrm{d}z) + d\gamma_{k+1}^2 \right\} \; . \tag{17}$$

In addition, by the Pinsker inequality, $\|\delta_y Q_\gamma^{n+1,p} - \delta_y P^{\mathrm{L}}_{\Gamma_{n+1},p}\|_{\mathrm{TV}} \le \sqrt{2}\{\mathrm{KL}(\mu^y_{n,p}|\bar\mu^y_{n,p})\}^{1/2}$. So, combining (3) and (17) in (13), we have for all $x \in \mathbb{R}^d$, $n \ge 0$, $p \ge 1$, $n < p$,

$$\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \le 2^{-1/2} L \left(\sum_{k=n}^{p-1} \left\{ (\gamma_{k+1}^3/3) A(\gamma, x) + d\gamma_{k+1}^2 \right\} \right)^{1/2} + C(\delta_x Q_\gamma^n) \kappa^{\Gamma_{n+1,p}} \; , \tag{18}$$

where $\kappa, C(\delta_x Q_\gamma^n)$ are defined in (3) and

$$A(\gamma, x) = \sup_{k\ge 0} \int_{\mathbb{R}^d} \|\nabla U(z)\|^2 \, Q_\gamma^k(y, \mathrm{d}z) \; . \tag{19}$$

In the sequel, depending on the assumption and the technique of proof, for any given $x \in \mathbb{R}^d$, $C(\delta_x Q_\gamma^n)$ can have two kinds a upper bound, either of the form $-\log(\gamma_n) W(x)$, or $\exp(a\Gamma_n) W(x)$,

for some function $W : \mathbb{R}^d \to \mathbb{R}$ and $a > 0$. In both cases, as shown in Proposition 2, it is possible to choose n as a function of p, so that $\lim_{p\to+\infty} \|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} = 0$ under appropriate conditions on the sequence of step sizes $(\gamma_k)_{k\geq 1}$.

Proposition 2. *Assume* **L***1. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence such that $\lim_{k\to+\infty} \Gamma_k = +\infty$ and $\lim_{k\to\infty} \gamma_k = 0$. Then, $\lim_{n\to\infty} \|\delta_x Q_\gamma^n - \pi\|_{\mathrm{TV}} = 0$ for any $x \in \mathbb{R}^d$ for which one of the two following conditions holds:*

(i) $A(\gamma, x) < \infty$ and $\limsup_{n\to+\infty} C(\delta_x Q_\gamma^n)/(-\log(\gamma_n)) < +\infty$.
(ii) $\sum_{k=1}^{\infty} \gamma_k^2 < +\infty$, $A(\gamma, x) < \infty$ and $\limsup_{n\to+\infty} \log\{C(\delta_x Q_\gamma^n)\}/\Gamma_n < +\infty$.

Proof. (i) There exists $p_0 \geq 1$ such that for all $p \geq p_0$, $\kappa^{\gamma_p} > \gamma_p$ and $\kappa^{\Gamma_p} \leq \gamma_1$. Therefore, we can define for all $p \geq p_0$,

$$n(p) \stackrel{\mathrm{def}}{=} \min\left\{k \in \{0, \cdots, p-1\} \,|\kappa^{\Gamma_{k+1,p}} > \gamma_{k+1}\right\} . \tag{20}$$

and $n(p) \geq 1$. We first show that $\liminf_{p\to\infty} n(p) = \infty$. The proof goes by contradiction. If $\liminf_{p\to\infty} n(p) < \infty$ we could extract a bounded subsequence $(n(p_k))_{k\geq 1}$. For such sequence, $(\gamma_{n(p_k)+1})_{k\geq 1}$ is bounded away from 0, but $\lim_{k\to+\infty} \kappa^{\Gamma_{n(p_k)+1,p_k}} = 0$ which yields to a contradiction. The definition of $n(p)$ implies that $\kappa^{\Gamma_{n(p),p}} \leq \gamma_{n(p)}$, showing that

$$\limsup_{p\to+\infty} C(\delta_x Q_\gamma^{n(p)})\kappa^{\Gamma_{n(p),p}} \leq \limsup_{p\to+\infty} \frac{C(\delta_x Q_\gamma^{n(p)})}{-\log(\gamma_{n(p)})} \limsup_{p\to+\infty} \left\{\gamma_{n(p)}(-\log(\gamma_{n(p)}))\right\} = 0 .$$

On the other hand, since $(\gamma_k)_{k\geq 1}$ is nonincreasing, for any $\ell \geq 2$,

$$\sum_{k=n(p)+1}^{p} \gamma_k^\ell \leq \gamma_{n(p)+1}^{\ell-1} \Gamma_{n(p)+1,p} \leq \gamma_{n(p)+1}^{\ell-1} \log(\gamma_{n(p)+1})/\log(\kappa) .$$

The proof follows from (18) using $\lim_{p\to\infty} \gamma_{n(p)} = 0$.

(ii) For all $p \geq 1$, define $n(p) = \max(0, \lfloor \log(\Gamma_p) \rfloor)$. Note that $\lim_{k\to+\infty} \Gamma_k = +\infty$ implies that $\lim_{p\to+\infty} n(p) = +\infty$. Since by assumption, $\sum_{k=1}^{+\infty} \gamma_k^2 < +\infty$ and $(\gamma_k)_{k\geq 1}$ is a nonincreasing sequence, we get for all $\ell \geq 2$, $\lim_{p\to+\infty} \sum_{k=n(p)}^{p} \gamma_k^\ell = 0$, which shows that the first term in the right side of (18) goes to 0 as p goes to infinity. As for the second term, under the condition $\limsup_{n\to+\infty} \log\{C(\delta_x Q_\gamma^n)\}/\Gamma_n < +\infty$, we get using that $(\gamma_k)_{k\geq 1}$ is nonincreasing and $n(p) \leq \log(\Gamma_p)$,

$$\begin{aligned} C(\delta_x Q_\gamma^{n(p)})\kappa^{\Gamma_{n(p),p}} &\leq \exp\left(\log(\kappa)\Gamma_p + \left[\{\log(C(\delta_x Q_\gamma^{n(p)}))/\Gamma_{n(p)}\}_+ - \log(\kappa)\right]\Gamma_{n(p)}\right) \\ &\leq \exp\left(\log(\kappa)\Gamma_p + \left[\sup_{k\geq 1}\{\log(C(\delta_x Q_\gamma^k))/\Gamma_k\}_+ - \log(\kappa)\right]\gamma_1 \log(\Gamma_p)\right) . \end{aligned}$$

Using $\kappa < 1$ and again $\lim_{k\to+\infty} \Gamma_k = +\infty$, we have $\lim_{p\to+\infty} C(\delta_x Q_\gamma^{n(p)})\kappa^{\Gamma_{n(p),p}} = 0$, which concludes the proof.

□

Using (18), we can also assess the convergence of the algorithm for constant step sizes $\gamma_k = \gamma$ for all $k \geq 1$. Two different kinds of results can be derived. First, for a given precision $\varepsilon > 0$, we can try to optimize the step size γ to minimize the number of iterations p required to achieve $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$. Second if the total number of iterations is fixed $p \geq 1$, we may determine the stepsize $\gamma > 0$ which minimizes $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}}$.

Lemma 3. *Assume that there exists $\bar{\gamma} > 0$ such that for any $\gamma \in (0, \bar{\gamma}]$ (18) holds for $\gamma_k = \gamma$ for all $k \geq 1$. In addition, assume that $\bar{C}(x) = \sup_{\gamma \in (0,\bar{\gamma}]} \sup_{n \geq 1} C(\delta_x Q_\gamma^n) < +\infty$ and $\sup_{\gamma \in (0,\bar{\gamma}]} A(\gamma, x) \leq \bar{A}(x)$. Then for all $\varepsilon > 0$, we get $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$ if*

$$p > T\gamma^{-1} \quad \text{and} \quad \gamma = \frac{-d + \sqrt{d^2 + (2/3)\bar{A}(x)\varepsilon^2 (L^2 T)^{-1}}}{2\bar{A}(x)/3} \wedge \bar{\gamma} \,, \tag{21}$$

where

$$T = \left(\sup_{n \geq 0} \log\{\bar{C}(x)\} - \log(\varepsilon/2) \right) \Big/ (-\log(\kappa)) \,.$$

Proof. For $p > T\gamma^{-1}$, set $n = p - \lfloor T\gamma^{-1} \rfloor$. Then using the stated expressions of γ and T in (18) concludes the proof. □

Lemma 4. *Assume that there exists $\bar{\gamma} > 0$ such that for any $\gamma \in (0, \bar{\gamma}]$ (18) holds for $\gamma_k = \gamma$ for all $k \geq 1$. In addition, assume that there exists $n > 0$ such that $\bar{C}_n(x) = \sup_{\gamma \in (0,\bar{\gamma}]} C(\delta_x Q_\gamma^n) < +\infty$ and $\sup_{\gamma \in (0,\bar{\gamma}]} A(\gamma, x) \leq \bar{A}(x)$. For all $p > n$ and all $x \in \mathbb{R}^d$, if $\gamma = \log(p-n)\{(p-n)(-\log(\kappa))\}^{-1} \leq \bar{\gamma}$, then*

$$\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq (p-n)^{-1/2} \{\bar{C}_n(x)(p-n)^{-1/2} + \log(p-n)(d + \bar{A}(x)\log(p-n)(p-n)^{-1})^{1/2}\} \,.$$

Proof. The proof is a straightforward calculation using (18). □

To get quantitative bounds for the total variation distance $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}}$ it is therefore required to get bounds on κ, $A(\gamma, x)$ and to control $C(\delta_x Q_\gamma^n)$. We will consider in the sequel two different approaches, one based on functional inequalities, the other on coupling techniques. We will consider also increasingly stringent assumptions for the potential U. Whereas we will qualitatively obtain always the same type of exponential bounds, the dependence of the constants in the dimension will be markedly different. In the worst case, the dependence is exponential. It is polynomial when U is convex.

2.1. Superexponential densities

The first condition that we consider is that the potential is superexponential outside a ball. This is a rather weak assumption (we do not assume convexity here). The price to pay will be constants which are exponential in the dimension. In this section, the potential U is unbounded off compact set. Since U is continuous, it has a global minimizer $x^\star$, which is a point at which $\nabla U(x^\star) = 0$. Without loss of generality, it is assumed that $U(x^\star) = 0$.

H 1. *The potential U is twice continuously differentiable and there exist $\rho > 0$, $\alpha \in (1, 2]$ and $M_\rho \geq 0$ such that for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq M_\rho$, $\langle \nabla U(x), x - x^\star \rangle \geq \rho \|x - x^\star\|^\alpha$.*

Lemma 5. *Assume **L** 1 and **H** 1. Then for all $x \in \mathbb{R}^d$,*

$$U(x) \geq \rho \|x - x^\star\|^\alpha / (\alpha + 1) - a_\alpha \quad \text{with} \quad a_\alpha = \rho M_\rho^\alpha / (\alpha + 1) + M_\rho^2 L/2 \,. \tag{22}$$

Proof. The elementary proof is postponed to Section 3.2. □

Following [31, Theorem 2.3], we first establish a drift condition for the diffusion.

Proposition 6. *Assume* ***L****1 and* ***H****1. For any* $\varsigma \in (0,1)$*, the drift condition* (5) *is satisfied with the Lyapunov function* $V_\varsigma(x) = \exp(\varsigma U(x))$*,* $\theta_\varsigma = \varsigma dL$*,* $\mathcal{E}_\varsigma = \mathrm{B}(x^\star, K_\varsigma)$*,* $K_\varsigma = \max(\{2dL/(\rho(1-\varsigma))\}^{1/(2(\alpha-1))}, M_\rho)$ *and* $\beta_\varsigma = \varsigma dL \sup_{\{y\in\mathcal{E}_\varsigma\}}\{V_\varsigma(y)\}$*. Moreover, there exists constant* $C_\varsigma < \infty$ *and* $\upsilon_\varsigma > 0$ *such that for all probability measures* μ_0 *and* ν_0 *on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*, satisfying* $\mu_0(V_\varsigma) + \nu_0(V_\varsigma) < +\infty$*, and* $t \in \mathbb{R}_+$*,*

$$\|\mu_0 P_t^{\mathrm{L}} - \nu_0 P_t^{\mathrm{L}}\|_{V_\varsigma} \le C_\varsigma \mathrm{e}^{-\upsilon_\varsigma t}\|\mu_0 - \nu_0\|_{V_\varsigma}\ ,\ \|\mu_0 P_t^{\mathrm{L}} - \pi\|_{V_\varsigma} \le C_\varsigma \mathrm{e}^{-\upsilon_\varsigma t}\mu_0(V_\varsigma)\ .$$

Proof. The proof, adapted from [31, Theorem 2.3] and [27, Theorem 6.1], is postponed to Section 3.3. □

The constants C_ς and υ_ς depend on the drift and minorization conditions. Under **H**1, explicit expressions for these constants have been developed in the literature but these estimates are in general very conservative. We now turn to establish a Foster-Lyapunov condition for the Euler discretization.

Proposition 7. *Assume* ***L****1 and* ***H****1. Let* $\bar\gamma \in \left(0, L^{-1}\right)$*. For all* $\gamma \in (0, \bar\gamma]$ *and* $x \in \mathbb{R}^d$*,* R_γ *satisfies the drift condition* (9) *with* $V(x) = \exp(U(x)/2)$*,* $\lambda = \mathrm{e}^{-dL/\{2(1-L\bar\gamma)\}}$*,* $c = -2\log(\lambda)\lambda^{-\bar\gamma}\sup_{\{y\in\mathrm{B}(x^\star,K)\}} V(y)$ *and* $K = \max(M_\rho, (8\log(\lambda)/\rho^2)^{1/(2(\alpha-1))})$*.*

Proof. The proof is postponed to Section 3.4. □

Theorem 8. *Assume* ***L****1 and* ***H****1. Let* $(\gamma_k)_{k\ge1}$ *be a nonincreasing sequence with* $\gamma_1 < \bar\gamma$*,* $\bar\gamma \in \left(0, L^{-1}\right)$*. Then, for all* $n \ge 0$*,* $p \ge 1$*,* $n < p$*, and* $x \in \mathbb{R}^d$*,* (18) *holds with* $\log(\kappa) = -\upsilon_{1/2}$ *and*

$$A(\gamma, x) \le L^2\left(\frac{\alpha+1}{\rho}\left[a_\alpha + \frac{4(2-\alpha)(\alpha+1)}{\alpha\rho} + 2\log\left\{G(\lambda, c, \gamma_1, V(x))\right\}\right]\right)^{2/\alpha} \tag{23}$$

$$C(\delta_x Q_\gamma^n) \le C_{1/2} F(\lambda, \Gamma_{1,n}, c, \gamma_1, V(x))\ , \tag{24}$$

where $C_{1/2}$*,* $\upsilon_{1/2}$ *are given by Proposition 6,* F *by* (11)*,* V*,* λ*,* c *in Proposition 7,* G *by* (12)*,* a_α *in* (22)*.*

Proof. The proof is postponed to Section 3.5. □

Equation (24) implies that for all $x \in \mathbb{R}^d$, we have $\sup_{n\ge0} C(\delta_x Q_\gamma^n) \le G(\lambda, c, \gamma_1, V(x))$, so Proposition 2-(i) shows that $\lim_{p\to+\infty}\|Q_\gamma^p - \pi\|_{\mathrm{TV}} = 0$, if $\lim_{k\to+\infty}\gamma_k = 0$ and $\lim_{k\to+\infty}\Gamma_k = +\infty$. In addition, for the case of constant step size $\gamma_k = \gamma$ for all $k \ge 1$, Lemma 3 and Lemma 4 can be applied.

By Proposition 6, $(P_t^{\mathrm{L}})_{t\ge0}$ is a contraction operator on the space of bounded signed measure $\mu \in \mathbb{M}_0$, $\mu(V_{1/4}) < +\infty$, endowed with $\|\cdot\|_{V_{1/4}}$. It is therefore possible to control $\|\delta_x Q_\gamma^p - \pi\|_{V_{1/4}}$. To simplify the notations, we limit our discussion to constant step sizes.

Theorem 9. *Assume* ***L****1 and* ***H****1. Then, for all* $p \ge 1$*,* $x \in \mathbb{R}^d$ *and* $\gamma \in \left(0, L^{-1}\right)$*, we have*

$$\|\delta_x R_\gamma^p - \pi\|_{V^{1/2}} \le C_{1/4}\kappa^{\gamma p} V^{1/2}(x) + B(V(x), \gamma)\ , \tag{25}$$

where $\log(\kappa) = -\upsilon_{1/4}$*,* $C_{1/4}, \upsilon_{1/4}, \theta_{1/2}, \beta_{1/2}$ *are defined in Proposition 6,* V, λ, c *in Proposition 7,* G *in* (12) *and*

$$B^2(v,\gamma) = L^2\max(1, C_{1/4}^2)(1+\gamma)(1-\kappa)^{-2}\left(2G(\lambda, c, \gamma, v) + \beta_{1/2}/\theta_{1/2}\right) \times \left(\gamma d + 3^{-1}\gamma^2\|\nabla U\|_{V^{1/2}}^2 G(\lambda, c, \gamma, v)\right)\ . \tag{26}$$

Moreover, R_γ *has a unique invariant distribution* π_γ *and* $\|\pi - \pi_\gamma\|_{V^{1/2}} \le B(1, \gamma)$*.*

Proof. The proof of (25) is postponed to Section 3.6. The bound for $\|\pi - \pi_\gamma\|_{V^{1/2}}$ is an easy consequence of (25). By Proposition 13 and [26, Theorem 16.0.1], R_γ is $V^{1/2}$-uniformly ergodic: $\lim_{p\to+\infty} \|\delta_x R_\gamma^p - \pi_\gamma\|_{V^{1/2}} = 0$ for all $x \in \mathbb{R}^d$. (25) shows that for all $x \in \mathbb{R}^d$,

$$\|\pi - \pi_\gamma\|_{V^{1/2}} \le \lim_{p\to+\infty} \left\{ \|\delta_x R_\gamma^p - \pi\|_{V^{1/2}} + \|\delta_x R_\gamma^p - \pi_\gamma\|_{V^{1/2}} \right\} \le B(V(x), \gamma) \; .$$

Taking the minimum over $x \in \mathbb{R}^d$ concludes the proof. □

Remark 10. *It is shown in [33, Theorem 4] that for $\phi \in C^\infty(\mathbb{R}^d)$ with polynomial growth, $\pi_\gamma(\phi) - \pi(\phi) = b(\phi)\gamma + \mathcal{O}(\gamma^2)$, for some constant $b(\phi) \ge 0$, provided that $U \in C^\infty(\mathbb{R}^d)$. Our result does not match this bound since $B(1,\gamma) = \mathcal{O}(\gamma^{1/2})$, but holds for a larger class of function ϕ, which can be only measurable. Our result also strengthens and completes [25, Corollary 7.5], which states that under **H1** with $\alpha = 2$, for any measurable $\phi : \mathbb{R}^d \to \mathbb{R}$, belonging to the function space*

$$\left\{ \phi : \mathbb{R}^d \to \mathbb{R} \; | \forall x, y \in \mathbb{R}^d \, , |\phi(x) - \phi(y)| \le C \, \|x - y\| \left\{ 1 + \|x\|^k + \|y\|^k \right\} \right\} \; ,$$

for some $C \ge 0$, $k \ge 1$, $|\pi_\gamma(\phi) - \pi(\phi)| \le C\gamma^\chi$ for some constants $C \ge 0$ and $\chi \in (0, 1/2)$, which does not depend on ϕ.

The bounds in Theorem 8 and Theorem 9 depend upon the constants appearing in Proposition 6 which are computable but are are known to be extremely pessimistic in general; see [32]. More explicit rates of convergence for the semigroup can be obtained using Poincaré (or spectral gap); see [2], [7] and [3, Chapter 4] and the references therein. The probability measure π is said to satisfy a Poincaré inequality with the constant C_{P} if for every locally Lipschitz function h,

$$\mathrm{Var}_\pi \{h\} \le C_{\mathrm{P}} \int_{\mathbb{R}^d} \|\nabla h(x)\|^2 \, \pi(\mathrm{d}x) \; . \tag{27}$$

This inequality implies by [7, Theorem 2.1] that for any initial distribution μ_0 such that $\mu_0 \ll \pi$ for all $t \ge 0$,

$$\left\| \mu_0 P_t^{\mathrm{L}} - \pi \right\|_{\mathrm{TV}} \le \exp(-t/C_{\mathrm{P}}) \left(\mathrm{Var}_\pi \{ \mathrm{d}\mu_0/\mathrm{d}\pi \} \right)^{1/2} \; . \tag{28}$$

[1, Theorem 1.4] shows that if the Lyapunov condition (5) is satisfied, then the Poincaré inequality (27) holds with constant depending explicitly on the Foster-Lyapunov constants and the oscillation of the function U in the ball $\mathrm{B}(0,K)$ defined by $\mathrm{osc}_K(U) = \sup_{\mathrm{B}(0,K)}(U) - \inf_{\mathrm{B}(0,K)}(U)$. Denote by

$$D_n(\gamma) \stackrel{\mathrm{def}}{=} \left(4\pi \left\{ \prod_{k=1}^n (1 - L\gamma_k) \right\}^2 \sum_{i=1}^n \gamma_i (1 - L\gamma_i)^{-1} \right)^{-d/2} \; . \tag{29}$$

Theorem 11. *Assume **L1** and **H1**. Let $(\gamma_k)_{k\ge1}$ be a non increasing sequence. Then for all $n \ge 1$ and $x \in \mathbb{R}^d$, Equation (3) holds with*

$$\log(\kappa) = \left(-\theta_{1/2}^{-1} \left\{ 1 + (4\beta_{1/2} K_{1/2}^2/\pi^2) \mathrm{e}^{\mathrm{osc}_{K_{1/2}}(U)} \right\} \right)^{-1} \; ,$$

$$C(\delta_x Q_\gamma^n) \le \frac{(\alpha+1)^d (2\pi)^{(d+1)/2} (d-1)!}{\rho^d \mathbf{\Gamma}((d+1)/2)} D_n(\gamma) \mathrm{e}^{a_\alpha} \, \mathrm{e}^{U(x)} \; ,$$

where $\mathbf{\Gamma}$ is the Gamma function and the constants $\beta_{1/2}, \theta_{1/2}, K_{1/2}, a_\alpha$ are given in Proposition 6 and (22) respectively.

Proof. The proof consists in giving a bound on $\mathrm{Var}_\pi\{\mathrm{d}\delta_x Q_\gamma^n/\mathrm{d}\pi\}$ and is postponed to Section 3.7. □

Note that for all $x \in \mathbb{R}^d$, $C(\delta_x Q_\gamma^n)$ satisfies the conditions of Proposition 2-(ii). Therefore, if $\lim_{n\to+\infty} \Gamma_n = +\infty$ and $\lim_{n\to+\infty} \sum_{k=1}^n \gamma_k^2 < +\infty$, we get $\lim_{k\to+\infty} \|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} = 0$.

2.2. Log-concave densities

We now derive non-asymptotic bounds under the following additional assumption.

H2. *U is convex and there exists a minimizer $x^\star$ for U and for some $\eta > 0$ and $M_\eta \geq 0$, for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq M_\eta$,*

$$U(x) - U(x^\star) \geq \eta \|x - x^\star\| \ . \tag{30}$$

It is shown in [1, Lemma 2.2] that if U satisfies **L**1 and is convex, then (30) holds for some constants η, M_η which depend in an intricate way on U. Since the constants η, M_η appear explicitly in the bounds that we derive, we must assume that these constants are explicitly computable. We still assume in this section that $U(x^\star) = 0$. Define the function $W_{\mathrm{c}} : \mathbb{R}^d \to [1, +\infty)$ for all $x \in \mathbb{R}^d$ by

$$W_{\mathrm{c}}(x) = \exp((\eta/4)(\|x - x^\star\|^2 + 1)^{1/2}) \ . \tag{31}$$

Proposition 12. *Assume **L**1 and **H**2. Then the drift condition (5) is satisfied with the Lyapunov function W_{c}, $\theta = \eta^2/8$, $\mathcal{E} = \mathrm{B}(x^\star, K)$, $K = \max(1, M_\eta, 4d/\eta)$ and*

$$\beta = (\eta/4)\left[\eta/4 + d + \sup_{y \in \mathrm{B}(x^\star, K)} \{\|\nabla U(y)\|\}\right] \max\left\{1, (K^2+1)^{-1/2} \exp(\eta(K^2+1)^{1/2}/4)\right\} \ .$$

Proof. The proof is adapted from [1, Corollary 1.6] and is postponed to Section 3.8. □

We now derive a drift inequality for R_γ in the case of U is convex.

Proposition 13. *Assume **L**1 and **H**2. Let $\bar\gamma \in \left(0, L^{-1}\right]$. Then for all $\gamma \in (0, \bar\gamma]$, R_γ satisfies (9) with the Lyapunov function W_{c}, $\lambda = \mathrm{e}^{-2^{-4}\eta^2(2^{1/2}-1)}$, $R_{\mathrm{c}} = \max(1, 2d/\eta, M_\eta)$,*

$$c = \{(\eta/4)(d + (\eta\bar\gamma/4)) - \log(\lambda)\}\, \mathrm{e}^{\eta(R_{\mathrm{c}}^2+1)^{1/2}/4 + (\eta\bar\gamma/4)(d+(\eta\bar\gamma/4))} \ . \tag{32}$$

Proof. The proof is postponed to Section 3.9 □

If U is convex, [4, Theorem 1.2] shows that π satisfies a Poincaré inequality with a constant depending only on the variance of the stationary distribution.

Theorem 14. *Assume **L** 1 and **H** 2. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq \bar\gamma$, $\bar\gamma \in \left(0, L^{-1}\right)$. Then, for all $n \geq 0$, $p \geq 1$, $n < p$, and $x \in \mathbb{R}^d$, (18) holds with*

$$\log(\kappa) = \left(-432 \int_{\mathbb{R}^d} \left\| x - \int_{\mathbb{R}^d} y \pi(\mathrm{d}y) \right\|^2 \pi(\mathrm{d}x)\right)^{-1} \tag{33a}$$

$$C(\delta_x Q_\gamma^n) = \left(\frac{(2\pi)^{(d+1)/2}(d-1)!}{\eta^d \boldsymbol{\Gamma}((d+1)/2)} + \frac{\pi^{d/2} M_\eta^d}{\boldsymbol{\Gamma}(d/2+1)}\right) D_n(\gamma) \exp(U(x)) \tag{33b}$$

$$A(\gamma, x) = \left(4\eta^{-1}\left[1 + \log\left\{G(\lambda, c, \gamma_1, W_{\mathrm{c}}(x))\right\}\right]\right)^2 \ , \tag{33c}$$

where $D_n(\gamma)$, G, W_{c}, λ, c, are given in (29), (12), (31), Proposition 13 respectively.

Proof. The proof is postponed to Section 3.10. □

For all $x \in \mathbb{R}^d$, $C(\delta_x Q_\gamma^n)$ satisfies the conditions of Proposition 2-(ii). Therefore, if $\lim_{n\to+\infty} \Gamma_n = +\infty$ and $\lim_{n\to+\infty} \sum_{k=1}^n \gamma_k^2 < +\infty$, we get $\lim_{k\to+\infty} \|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} = 0$.

There are two difficulties when applying Theorem 14. First the Poincaré constant (33a) is in closed form but is not computable, although it can be bounded by a $\mathcal{O}(d^2)$. Second, the control of

	d	ε	L
γ	$\mathcal{O}(d^{-4})$	$\mathcal{O}(\varepsilon^2/\log(\varepsilon^{-1}))$	$\mathcal{O}(L^{-2})$
p	$\mathcal{O}(d^7)$	$\mathcal{O}(\varepsilon^{-2}\log^2(\varepsilon^{-1}))$	$\mathcal{O}(L^2)$

TABLE 1
For constant step sizes, dependence of γ and p in d, ε and parameters of U to get $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$ using Theorem 17

the variance $\mathrm{Var}_\pi\{\mathrm{d}\delta_x Q_\gamma^n/\mathrm{d}\pi\}$ is difficult and yields to bounds which are likely to be suboptimal yet difficult to improve. To circumvent this two issues, we now give new quantitative results on the convergence of $(P_t^{\mathrm{L}})_{t\geq 0}$ to π in total variation. Instead of using functional inequality, we use in the proof the coupling by reflection, introduced in [24]. Define the function $\omega : (0,1)\times\mathbb{R}_+^* \to \mathbb{R}_+$ by for all $\epsilon \in (0,1)$ and $R \geq 0$,

$$\omega(\epsilon, R) = R^2 / \left\{2\boldsymbol{\Phi}^{-1}(1-\epsilon/2)\right\}^2 , \tag{34}$$

where $\boldsymbol{\Phi}$ is the cumulative distribution function of the standard Gaussian and $\boldsymbol{\Phi}^{-1}$ is the associated quantile function.

Theorem 15. *Assume **L**1 and **H**2. Then for all $x \in \mathbb{R}^d$, Equation (3) is satisfied with*

$$\log(\kappa) = -2\log(2)\theta \left[\log\left\{\beta\left(2 + 2\theta^{-1}\mathrm{e}^{2\theta^{-1}\omega\left(2^{-1},(8/\eta)\log(4\theta^{-1}\beta)\right)}\right)\right\} + \log(2)\right]^{-1} , \tag{35a}$$

$$C(\delta_x Q_\gamma^n) \leq 4 + (1/2)\left\{\beta\theta^{-1} + F(\lambda, \Gamma_{1,n}, c, \gamma_1, W_{\mathrm{c}}(x))\right\} , \tag{35b}$$

where the functions F and W_{c} are defined in (11) and (31), the constants $\theta, \beta, \lambda, c$ in Proposition 12 and Proposition 13 respectively.

Proof. The proof is postponed to Section 4.1. □

Contrary to (33b), (35b) is uniformly bounded in n and we can apply Proposition 2-(i), which implies the convergence to 0 of $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}}$ as p goes to infinity, if $\lim_{k\to+\infty}\gamma_k = 0$ and $\lim_{k\to+\infty}\Gamma_k = +\infty$. On the other hand, the expression of the rate (35a) is less concise and explicit than (33a). Nevertheless, since $\log(\beta)$ in Proposition 12 is of order d, we get that the rate of convergence $\log(\kappa)$ is of order d^{-2} as d goes to infinity (note indeed that the leading term when d is large is $2\theta^{-1}\omega\left(2^{-1}, (8/\eta)\log(4\theta^{-1}\beta)\right)$ which is of order d^2). In the case of constant step sizes $\gamma_k = \gamma$ for all $k \geq 0$, we can apply both Lemma 3 and Lemma 4. In particular, we summarise in Table 1 the dependence of the step size $\gamma > 0$ and the minimum number of iterations $p \geq 0$, given by Lemma 3, to achieve $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$ for $\varepsilon > 0$.

2.3. Strongly log-concave densities

More precise bounds can be obtained in the case where U is assumed to be strongly convex outside some ball; this assumption has been considered by [12] for convergence in the Wasserstein distance; see also [5].

H3 (M_{s})**.** *U is convex and there exist $M_{\mathrm{s}} \geq 0$ and $m > 0$, such that for all $x, y \in \mathbb{R}^d$, $\|x - y\| \geq M_{\mathrm{s}}$,*

$$\langle \nabla U(x) - \nabla U(y), x - y\rangle \geq m\|x - y\|^2 .$$

We will see in the sequel that under this assumption the convergence rate does not depend on the dimension d but only on the constants m and M_{s}.

	d	ε	L	m	M_{s}
γ	$\mathcal{O}(d^{-1})$	$\mathcal{O}(\varepsilon^2/\log(\varepsilon^{-1}))$	$\mathcal{O}(L^{-2})$	$\mathcal{O}(m)$	$\mathcal{O}(M_{\mathrm{s}}^{-4})$
p	$\mathcal{O}(d\log(d))$	$\mathcal{O}(\varepsilon^{-2}\log^2(\varepsilon^{-1}))$	$\mathcal{O}(L^2)$	$\mathcal{O}(m^{-2})$	$\mathcal{O}(M_{\mathrm{s}}^{8})$

TABLE 2
For constant step sizes, dependence of γ and p in d, ε and parameters of U to get $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$ using Theorem 17

Proposition 16. *Assume **L**1 and **H**3(M_{s}). Let $\bar\gamma \in \left(0, 2mL^{-2}\right)$. For all $\gamma \in (0, \bar\gamma]$, the Markov kernel R_γ satisfies the Foster-Lyapunov condition (9) with $V(x) = \|x - x^\star\|^2$, $\lambda = \mathrm{e}^{-2m+\bar\gamma L^2}$ and $\mathrm{c} = 2(d + mM_{\mathrm{s}}^2)$.*

Proof. The proof is postponed to Section 3.11. □

Theorem 17. *Assume **L**1 and **H**3(M_{s}). Let $(\gamma_k)_{k\in\mathbb{N}^*}$ be a nonincreasing sequence with $\gamma_1 \leq \bar\gamma$, $\bar\gamma \in \left(0, 2mL^{-2}\right)$. Then, for all $n \geq 0$, $p \geq 1$, $n < p$, and $x \in \mathbb{R}^d$, (18) holds with*

$$\log(\kappa) = -(m/2)\log(2)\left[\log\left\{\left(1 + \mathrm{e}^{m\omega(2^{-1},\max(1,M_{\mathrm{s}}))/4}\right)(1 + \max(1, M_{\mathrm{s}}))\right\} + \log(2)\right]^{-1}$$
$$C(\delta_x Q_\gamma^n) \leq 5 + \left(d/m + M_{\mathrm{s}}^2\right)^{1/2} + F^{1/2}(\lambda, \Gamma_{1,n}, \mathrm{c}, \gamma_1, \|x - x^\star\|^2)$$
$$A(\gamma, x) \leq L^2\, G(\lambda, \mathrm{c}, \gamma_1, \|x - x^\star\|^2)\ ,$$

where λ, c are given in Proposition 16.

Proof. The proof is postponed to Section 4.1. □

Note that the conditions of Proposition 2-(i) are fulfilled. For constant step sizes $\gamma_k = \gamma$ for all $k \geq 1$, Lemma 3 and Lemma 4 can be applied. We give in Table 2 the dependence of the step size $\gamma > 0$ and the minimum number of iterations $p \geq 0$, provided in Lemma 3, on the dimension d and the other constants related to U, to get $\|\delta_x Q_\gamma^p - \pi\|_{\mathrm{TV}} \leq \varepsilon$, for a target precision $\varepsilon > 0$. We can see that the dependence on the dimension is milder than for the convex case.

We now consider the case where U is a bounded perturbation of a strongly convex potential.

H4. *The potential U may be expressed as $U = U_1 + U_2$, where*

*(a) $U_1 : \mathbb{R}^d \to \mathbb{R}$ satisfies **H**3(0) (i.e. is strongly convex) and there exists $L_1 \geq 0$ such that for all $x, y \in \mathbb{R}^d$,*

$$\|\nabla U_1(x) - \nabla U_1(y)\| \leq L_1 \|x - y\|\ .$$

(b) $U_2 : \mathbb{R}^d \to \mathbb{R}$ is continuously differentiable and $\sup_{x\in\mathbb{R}^d}\{|U_2(x)| + \|\nabla U_2(x)\|\} < +\infty$.

The probability measure π is said to satisfy a log-Sobolev inequality with constant $C_{\mathrm{LS}} > 0$ if for all locally Lipschitz function $h : \mathbb{R}^d \to \mathbb{R}$, we have

$$\mathrm{Ent}_\pi\left(h^2\right) \leq 2C_{\mathrm{LS}} \int \|\nabla h\|^2\, \mathrm{d}\pi\ .$$

Then [7, Theorem 2.7] shows that for any probability measure $\mu_0 \ll \pi$ satisfying $\mathrm{d}\mu_0/\mathrm{d}\pi \log(\mathrm{d}\mu_0/\mathrm{d}\pi) \in \mathrm{L}^1(\pi)$, for all $t \geq 0$, we have

$$\|\mu_0 P_t^{\mathrm{L}} - \pi\|_{\mathrm{TV}} \leq \mathrm{e}^{-t/C_{\mathrm{LS}}}\left\{2\,\mathrm{Ent}_\pi\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)\right\}^{1/2}\ . \tag{36}$$

Under **H** 4, [3, Corollary 5.7.2] and the Holley-Stroock perturbation principle [16, p. 1184], π satisfies a log-Sobolev inequality with a constant which only depends on the strong convexity constant m of U_1 and the oscillation of U_2 on $\mathbb{R}^d$, $\sup_{x\in\mathbb{R}^d} U_2(x) - \inf_{x\in\mathbb{R}^d} U_2(x)$. By [29, Theorem 2.1.12, Theorem 2.1.9], **H**4-(b) implies that for all $x, y \in \mathbb{R}^d$:

$$\langle \nabla U_1(y) - \nabla U_1(x), y - x\rangle \geq (\varpi/2)\left\|y-x\right\|^2 + \frac{1}{m+L_1}\left\|\nabla U_1(y) - \nabla U_1(x)\right\|^2 , \tag{37}$$

where

$$\varpi = \frac{2mL_1}{m+L_1} . \tag{38}$$

Denote by $x_1^\star$ the minimizer of U_1.

Proposition 18. *Assume **H**4. Let $(\gamma_k)_{k\geq 1}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L_1)$. Then for all $p \geq 1$ and $x \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} \left\|y - x_1^\star\right\|^2 Q_\gamma^p(x, \mathrm{d}y) \leq \prod_{k=1}^{p}(1 - \varpi\gamma_k/2)\left\|x - x_1^\star\right\|^2 + 2\varpi^{-1}(2d + (\gamma_1 + 2\varpi^{-1})\left\|\nabla U_2\right\|_\infty^2) .$$

Proof. The proof is postponed to Section 3.12. □

Theorem 19. *Assume **L**1 and **H**4. Let $(\gamma_k)_{k\in\mathbb{N}^*}$ be a nonincreasing sequence with $\gamma_1 \leq 2/(m+L_1)$. Then, for all $n, p \geq 1$, $n < p$, and $x \in \mathbb{R}^d$, (18) holds with*

$$\begin{aligned}
&-\log(\kappa) = m\exp\{-\mathrm{osc}_{\mathbb{R}^d}(U_2)\} \\
&C^2(\delta_x Q_\gamma^n) \leq L_1 \mathrm{e}^{-\varpi\Gamma_n/2}\left\|x - x_1^\star\right\|^2 + L_1\gamma_n(\gamma_n + 2\varpi^{-1})\left\|\nabla U_2\right\|_\infty^2 + 2\mathrm{osc}_{\mathbb{R}^d}(U_2) \\
&+ 2L_1\varpi^{-1}(1 - \varpi\gamma_n)(2d + (\gamma_1 + 2\varpi^{-1})\left\|\nabla U_2\right\|_\infty^2) - d(1 + \log(2\gamma_n m) - 2L_1\gamma_n) \\
&A(\gamma, x) \leq 2L_1^2\left\{\left\|x_1^\star - x^\star\right\|^2 + 2\varpi^{-1}(2d + (\gamma_1 + 2\varpi^{-1})\left\|\nabla U_2\right\|_\infty^2)\right\} + 2\left\|\nabla U_2\right\|_\infty^2 ,
\end{aligned} \tag{39}$$

where ϖ is defined in (38).

Proof. The proof is postponed to Section 3.13. □

Note that by (39), $\sup_{n\geq 1}\{C(\delta_x Q_\gamma^n)/(-\log(\gamma_n))\} < +\infty$, therefore Proposition 2-(i) can be applied and $\lim_{p\to+\infty}\left\|\delta_\gamma Q_\gamma^p - \pi\right\|_{\mathrm{TV}} = 0$ if $\lim_{k\to+\infty}\gamma_k = 0$ and $\lim_{k\to+\infty}\Gamma_k = +\infty$.

3. Proofs

3.1. Proof of Lemma 1

By a straightforward induction, we get for all $n \geq 0$ and $x \in \mathbb{R}^d$,

$$Q_\gamma^n V(x) \leq \lambda^{\Gamma_{1,n}} V(x) + c\sum_{i=1}^{n}\gamma_i\lambda^{\Gamma_{i+1,n}} . \tag{40}$$

Note that for all $n \geq 1$, we have since $(\gamma_k)_{k\geq 1}$ is nonincreasing and for all $t \geq 0$, $\lambda^t = 1 + \int_0^t \lambda^s \log(\lambda) \mathrm{d}s$,

$$\begin{aligned}\sum_{i=1}^n \gamma_i \lambda^{\Gamma_{i+1,n}} &\leq \sum_{i=1}^n \gamma_i \prod_{j=i+1}^n (1 + \lambda^{\gamma_1} \log(\lambda)\gamma_j) \\ &\leq (-\lambda^{\gamma_1} \log(\lambda))^{-1} \sum_{i=1}^n \gamma_i \left\{ \prod_{j=i+1}^n (1 + \lambda^{\gamma_1} \log(\lambda)\gamma_j) - \prod_{j=i}^n (1 + \lambda^{\gamma_1} \log(\lambda)\gamma_j) \right\} \\ &\leq (-\lambda^{\gamma_1} \log(\lambda))^{-1} \, .\end{aligned}$$

The proof is then completed using this inequality in (40).

3.2. Proof of Lemma 5

By **L**1, **H**1, the Cauchy-Schwarz inequality and $\nabla U(x^\star) = 0$, for all $x \in \mathbb{R}^d$, $\|x\| \geq M_\rho$, we have

$$\begin{aligned}&U(x) - U(x^\star) = \int_0^1 \langle \nabla U(x^\star + t(x - x^\star)), x - x^\star \rangle \, \mathrm{d}t \\ &\geq \int_0^{\frac{M_\rho}{\|x - x^\star\|}} \langle \nabla U(x^\star + t(x - x^\star)), x - x^\star \rangle \, \mathrm{d}t + \int_{\frac{M_\rho}{\|x - x^\star\|}}^1 \langle \nabla U(x^\star + t(x - x^\star)), t(x - x^\star) \rangle \, \mathrm{d}t \\ &\geq -M_\rho^2 L/2 + \rho \, \|x - x^\star\|^\alpha \, (\alpha + 1)^{-1} \left\{ 1 - (M_\rho / \|x - x^\star\|)^{\alpha+1} \right\} \, ,\end{aligned}$$

On the other hand using again **L**1, the Cauchy-Schwarz inequality and $\nabla U(x^\star) = 0$, for all $x \in \mathrm{B}(x^\star, M_\rho)$,

$$U(x) - U(x^\star) = \int_0^1 \langle \nabla U(x^\star + t(x - x^\star)), x - x^\star \rangle \, \mathrm{d}t \geq -M_\rho^2 L/2 \, ,$$

which concludes the proof.

3.3. Proof of Proposition 6

For all $x \in \mathbb{R}^d$, we have

$$\mathscr{A}^{\mathrm{L}} V_\varsigma(x) = \varsigma(1 - \varsigma) \left\{ - \|\nabla U(x)\|^2 + (1 - \varsigma)^{-1} \Delta U(x) \right\} V_\varsigma(x)$$

If $\alpha > 1$, by the Cauchy-Schwarz inequality, under **L**1-**H**1 for all $x \in \mathbb{R}^d$, $\Delta U(x) \leq dL$ and $\|\nabla U(x)\| \geq \rho \, \|x - x^\star\|^{\alpha - 1}$ for $\|x - x^\star\| \geq M_\rho$. Then, for all $x \notin \mathcal{E}_\varsigma$,

$$\mathscr{A}^{\mathrm{L}} V_\varsigma(x) \leq \varsigma(1 - \varsigma) \left\{ -\rho \, \|x - x^\star\|^{2(\alpha-1)} + (1 - \varsigma)^{-1} dL \right\} V_\varsigma(x) \leq -\varsigma dL V_\varsigma(x) \, ,$$

and $\sup_{\{x \in \mathcal{E}_\varsigma\}} \mathscr{A}^{\mathrm{L}} V_\varsigma(x) \leq \varsigma dL \sup_{\{y \in \mathcal{E}_\varsigma\}} \{V_\varsigma(y)\}$.

3.4. Proof of Proposition 7

By **H**1, for all $x \notin \mathrm{B}(x^\star, M_\rho)$,

$$\|\nabla U(x)\| \geq \rho \|x - x^\star\|^{\alpha - 1} \ . \tag{41}$$

Since under **L**1, for all $x, y \in \mathbb{R}^d$, $U(y) \leq U(x) + \langle \nabla U(x), y - x \rangle + (L/2)\|y - x\|^2$, we have for all $\gamma \in (0, \bar{\gamma})$ and $x \in \mathbb{R}^d$,

$$\begin{aligned} R_\gamma V(x)/V(x) &= (4\pi\gamma)^{-d/2} \int_{\mathbb{R}^d} \exp\left(\{U(y) - U(x)\}/2 - (4\gamma)^{-1} \|y - x + \gamma \nabla U(x)\|^2 \right) \mathrm{d}y \\ &\leq (4\pi\gamma)^{-d/2} \int_{\mathbb{R}^d} \exp\left(-4^{-1}\gamma \|\nabla U(x)\|^2 - (4\gamma)^{-1}(1 - \gamma L) \|y - x\|^2 \right) \mathrm{d}y \\ &\leq (1 - \gamma L)^{-d/2} \exp(-4^{-1}\gamma \|\nabla U(x)\|^2) \ , \end{aligned} \tag{42}$$

where we used in the last line that $\gamma < L^{-1}$. Since $\log(1 - L\gamma) = -L \int_0^\gamma (1 - Lt)^{-1} \mathrm{d}t$, for all $\gamma \in (0, \bar{\gamma}]$, $\log(1 - L\gamma) \geq -L\gamma(1 - L\bar{\gamma})^{-1}$. Using this inequality, (42) becomes

$$R_\gamma V(x)/V(x) \leq \lambda^{-\gamma} \exp\left(-4^{-1}\gamma \|\nabla U(x)\|^2 \right) \ . \tag{43}$$

By (41), for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq K$, we have

$$R_\gamma V(x) \leq \lambda^\gamma V(x) \ . \tag{44}$$

Also by (43) and since for all $t \geq 0$, $\mathrm{e}^t - 1 \leq t\mathrm{e}^t$, we get for all $x \in \mathbb{R}^d$

$$R_\gamma V(x) - \lambda^\gamma V(x) \leq \lambda^\gamma (\lambda^{-2\gamma} - 1) V(x) \leq -2\gamma \log(\lambda) \lambda^{-\bar{\gamma}} V(x) \ .$$

The proof is completed combining the last inequality and (44).

3.5. Proof of Theorem 8

We first consider (23). By **L**1, we have

$$\mathbb{E}_x[\|\nabla U(X_k)\|^2] \leq L^2 \mathbb{E}_x[\|X_k - x^\star\|^2] \ . \tag{45}$$

Consider now the function $\phi_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ defined for all $t \geq 0$ by $\phi_\alpha(t) = \exp(A_\alpha (t + B_\alpha)^{\alpha/2})$ where $A_\alpha = \rho/(2(\alpha + 1))$ and $B_\alpha = \{(2 - \alpha)/(\alpha A_\alpha)\}^{2/\alpha}$. Since ϕ_α is convex on $\mathbb{R}_+$ and is invertible on $\mathbb{R}_+$, we get using the Jensen inequality and Lemma 5 for all $k \geq 0$:

$$\mathbb{E}_x[\|X_k - x^\star\|^2] \leq \phi_\alpha^{-1} \left(\mathbb{E}_x[\phi_\alpha \left(\|X_k - x^\star\|^2 \right)] \right) \leq \phi_\alpha^{-1} \left(\mathrm{e}^{a_\alpha/2 + B_\alpha^{\alpha/2}} \mathbb{E}_x[V(X_k)] \right) \ ,$$

where $V(x) = \exp(U(x)/2)$. Using that for all $t \geq 0$, $\phi_\alpha^{-1}(t) \leq (A_\alpha^{-1} \log(t))^{2/\alpha}$ and Lemma 1, we get

$$\sup_{k \geq 0} \mathbb{E}_x[\|X_k - x^\star\|^2] \leq \left(A_\alpha^{-1} \left[a_\alpha/2 + B_\alpha^{\alpha/2} + \log\left\{ G(\lambda, c(\gamma_1), V(x)) \right\} \right] \right)^{2/\alpha} \ .$$

Eq. (24) follows from Proposition 6, Proposition 7 and Lemma 1.

3.6. Proof of Theorem 9

Lemma 20. *Let μ and ν be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $V : \mathbb{R}^d \to [1, \infty)$ be a measurable function. Then*

$$\|\mu - \nu\|_V \leq \sqrt{2} \left\{\nu(V^2) + \mu(V^2)\right\}^{1/2} \mathrm{KL}^{1/2}(\mu|\nu) \, .$$

Proof. Without losing any generality, we assume that $\mu \ll \nu$. For all $t \in [0,1]$, $t\log(t) - t + 1 = \int_t^1 (u-t)u^{-1}\mathrm{d}u \geq 2^{-1}(1-t)^2$, and on $[1, +\infty)$, $t \mapsto 2(1+t)(t\log(t) - t + 1) - (1-t)^2$ is nonincreasing. Therefore, for all $t \geq 0$,

$$|1 - t| \leq (2(1+t)(t\log(t) - t + 1))^{1/2} \, . \tag{46}$$

Then, we have:

$$\begin{aligned}
\|\mu - \nu\|_V &= \sup_{f \in \mathbb{F}(\mathbb{R}^d), \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f(x) \mathrm{d}\mu(x) - \int_{\mathbb{R}^d} f(x) \mathrm{d}\nu(x) \right| \\
&= \sup_{f \in \mathbb{F}(\mathbb{R}^d), \|f\|_V \leq 1} \left| \int_{\mathbb{R}^d} f(x) \left\{ \frac{\mathrm{d}\bar{\mu}}{\mathrm{d}\mu} - 1 \right\} \mathrm{d}\nu(x) \right| \leq \int_{\mathbb{R}^d} V(x) \left| \frac{\mathrm{d}\bar{\mu}}{\mathrm{d}\mu} - 1 \right| \mathrm{d}\nu(x) \, .
\end{aligned}$$

Using (46) and the Cauchy-Schwarz inequality in the previous inequality concludes the proof. □

Proof of Theorem 9. First note that by the triangle inequality and Proposition 6, for all $p \geq 1$

$$\|\pi - \delta_x Q_\gamma^p\|_{V^{1/2}} \leq C_{1/4} \kappa^{p\gamma} V^{1/2}(x) + \|\delta_x P_{\Gamma_p}^{\mathrm{L}} - \delta_x Q_\gamma^p\|_{V^{1/2}} \, . \tag{47}$$

We now bound the second term of the right hand side. Let $\mathrm{k}_\gamma = \lceil \gamma^{-1} \rceil$ and q_γ and r_γ be respectively the quotient and the remainder of the Euclidean division of p by k_γ. The triangle inequality implies $\|\delta_x P_{\Gamma_p}^{\mathrm{L}} - \delta_x Q_\gamma^p\|_{V^{1/2}} \leq A + B$ with

$$\begin{aligned}
A &= \left\| \delta_x Q_\gamma^{(\mathrm{q}_\gamma - 1)\mathrm{k}_\gamma} P_{\Gamma_{(\mathrm{q}_\gamma - 1)\mathrm{k}_\gamma, p}}^{\mathrm{L}} - \delta_x Q_\gamma^{(\mathrm{q}_\gamma - 1)\mathrm{k}_\gamma} Q_\gamma^{(\mathrm{q}_\gamma - 1)\mathrm{k}_\gamma + 1, p} \right\|_{V^{1/2}} \\
B &= \sum_{i=1}^{\mathrm{q}_\gamma} \left\| \delta_x Q_\gamma^{(i-1)\mathrm{k}_\gamma} P_{\Gamma_{(i-1)\mathrm{k}_\gamma + 1, p}}^{\mathrm{L}} - \delta_x Q_\gamma^{i\mathrm{k}_\gamma} P_{\Gamma_{i\mathrm{k}_\gamma + 1, p}}^{\mathrm{L}} \right\|_{V^{1/2}} \, .
\end{aligned}$$

It follows from Proposition 6 and $\mathrm{k}_\gamma \geq \gamma^{-1}$ that

$$B \leq \sum_{i=1}^{\mathrm{q}_\gamma} C_{1/4} \kappa^{\mathrm{q}_\gamma - i} \left\| \delta_x Q_\gamma^{(i-1)\mathrm{k}_\gamma} P_{\Gamma_{(i-1)\mathrm{k}_\gamma + 1, i\mathrm{k}_\gamma}}^{\mathrm{L}} - \delta_x Q_\gamma^{i\mathrm{k}_\gamma} \right\|_{V^{1/2}} \, . \tag{48}$$

We now bound each term of the sum in the right hand side. For all initial distribution ν_0 on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $i, j \geq 1$, $i < j$, it follows from Lemma 20, [19, Theorem 4.1, Chapter 2] and (17):

$$\begin{aligned}
&\|\nu_0 Q_\gamma^{i,j} - \nu_0 P_{\Gamma_{i,j}}\|_{V^{1/2}} \leq 2^{-1/2} \left(\nu_0 Q_\gamma^{i,j}(V) + \nu_0 P_{\Gamma_{i,j}}(V)\right)^{1/2} \left\{\mathrm{KL}(\nu_0 Q_\gamma^{i,j} | \nu_0 P_{\Gamma_{i,j}})\right\}^{1/2} \\
&\leq 2^{-1/2} L \left(\nu_0 Q_\gamma^{i,j}(V) + \nu_0 P_{\Gamma_{i,j}}(V)\right)^{1/2} \\
&\qquad \times (j-i)^{1/2} \left(\gamma^2 d + (\gamma^3/3) \sup_{k \in \{i, \cdots, j\}} \nu_0 Q_\gamma^{i,k-1}(\|\nabla U\|^2) \right)^{1/2} \, .
\end{aligned}$$

Proposition 6 implies by the proof of [27, Theorem 6.1] that for all $y \in \mathbb{R}^d$ and $t \geq 0$: $P_t^{\mathrm{L}} V(y) \leq V(y) + \beta_{1/2}/\theta_{1/2}$. Then, using Proposition 7, Lemma 1 and $\mathrm{k}_\gamma \geq \gamma^{-1}$ in (48), we get

$$\sup_{i \in \{1,\cdots,\mathrm{q}_\gamma\}} \left\| \delta_x Q_\gamma^{(i-1)\mathrm{k}_\gamma} P^{\mathrm{L}}_{\Gamma_{(i-1)\mathrm{k}_\gamma+1, i\mathrm{k}_\gamma}} - \delta_x Q_\gamma^{i\mathrm{k}_\gamma} \right\|^2_{V^{1/2}}$$
$$\leq 2^{-1}(1+\gamma) L^2 \left\{ 2G(\lambda, c, V(x)) + \beta_{1/2}/\theta_{1/2} \right\} \left\{ \gamma d + 3^{-1}\gamma^2 \|\nabla U\|^2_{V^{1/2}} G(\lambda, c, V(x)) \right\} .$$

Finally, A can be bounded along the same lines. □

3.7. Proof of Theorem 11

Denote for $\gamma > 0$, $r_\gamma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the transition density of R_γ defined for $x, y \in \mathbb{R}^d$ by

$$r_\gamma(x, y) = (4\pi\gamma)^{-1} \exp(-(4\gamma)^{-1} \|y - x + \gamma \nabla U(x)\|^2) . \tag{49}$$

For all $n \geq 1$, we denote by $q_\gamma^n : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the transition density associated with Q_γ^n defined by induction by: for all $x, y \in \mathbb{R}^d$

$$q_\gamma^1(x, y) = r_{\gamma_1}(x, y) , \qquad q_\gamma^{n+1}(x, y) = \int_{\mathbb{R}^d} q_\gamma^n(x, z) r_{\gamma_{n+1}}(z, y) \mathrm{d}z \text{ for } n \geq 1 . \tag{50}$$

Lemma 21. *Assume* **L** 1. *Let* $(\gamma_k)_{k \geq 1}$ *be a nonincreasing sequence with* $\gamma_1 < L$. *Then for all* $n \geq 1$ *and* $x, y \in \mathbb{R}^d$,

$$q_\gamma^n(x, y) \leq \frac{\exp\left(2^{-1}(U(x) - U(y)) - (2\sigma_{\gamma,n})^{-1} \|y - x\|^2\right)}{(2\pi\sigma_{\gamma,n} \prod_{i=1}^n (1 - L\gamma_i))^{d/2}} ,$$

where $\sigma_{\gamma,n} = \sum_{i=1}^n 2\gamma_i (1 - L\gamma_i)^{-1}$.

Proof. Under **L** 1, we have for all $x, y \in \mathbb{R}^d$, $U(y) \leq U(x) + \langle \nabla U(x), y - x \rangle + (L/2) \|y - x\|^2$, which implies that for all $\gamma \in \left(0, L^{-1}\right)$

$$r_\gamma(x, y) \leq (4\pi\gamma)^{-d/2} \exp\left(2^{-1}(U(x) - U(y)) - (1 - L\gamma)(4\gamma)^{-1} \|y - x\|^2\right) . \tag{51}$$

Then, the proof of the claimed inequality is by induction. By (51), the inequality holds for $n = 1$. Now assume that it holds for $n \geq 1$. By induction hypothesis and (51) applied for $\gamma = \gamma_{n+1}$, we have

$$q_\gamma^{n+1}(x, y) \leq (4\pi\gamma_{n+1})^{-d/2} \left\{ 2\pi\sigma_{\gamma,n} \prod_{i=1}^n (1 - L\gamma_i) \right\}^{-d/2} \exp\left(2^{-1}(U(x) - U(y))\right)$$
$$\times \int_{\mathbb{R}^d} \exp\left(-(2\sigma_{\gamma,n})^{-1} \|z - x\|^2 - (1 - L\gamma_{n+1})(4\gamma_{n+1})^{-1} \|z - y\|^2\right) \mathrm{d}z$$
$$\leq (4\pi\gamma_{n+1})^{-d/2} \left\{ 2\pi\sigma_{\gamma,n} \prod_{i=1}^n (1 - L\gamma_i) \right\}^{-d/2} (2\pi)^{d/2} (\sigma_{\gamma,n}^{-1} + (1 - L\gamma_{n+1})/(2\gamma_{n+1}))^{-d/2}$$
$$\times \exp\left(2^{-1}(U(x) - U(y)) - (2\sigma_{\gamma,n+1})^{-1} \|y - x\|^2\right) .$$

Rearranging terms in the last inequality concludes the proof. □

Lemma 22. *Assume* ***L***1 *and* ***H***1. *Then* $\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\mathrm{d}y \leq \vartheta_U$ *where*

$$\vartheta_U \stackrel{def}{=} \mathrm{e}^{a_\alpha} \frac{(2\pi)^{(d+1)/2}(d-1)!}{\eta^d \mathbf{\Gamma}((d+1)/2)} , \tag{52}$$

where a_α *is given in* (22).

Proof. By Lemma 5, for all $x \in \mathbb{R}^d$, $U(x) \geq \rho \|x - x^\star\|/(\alpha+1) - a_\alpha$. Using the spherical coordinates, we get

$$\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\mathrm{d}y \leq \mathrm{e}^{a_\alpha} \left\{(2\pi)^{(d+1)/2}/\mathbf{\Gamma}((d+1)/2)\right\} \int_0^{+\infty} \mathrm{e}^{-\rho t/(\alpha+1)} t^{d-1}\mathrm{d}t .$$

Then the proof is concluded by a straightforward calculation. □

Corollary 23. *Assume* ***L***1 *and* ***H***1. *Let* $(\gamma_k)_{k\geq 1}$ *be a nonincreasing sequence with* $\gamma_1 < L$. *Then for all* $n \geq 1$ *and* $x \in \mathbb{R}^d$,

$$\mathrm{Var}_\pi \left\{\frac{\mathrm{d}\delta_x Q_\gamma^n}{\mathrm{d}\pi}\right\} \leq (\vartheta_U \exp(U(x))) \left(4\pi \left\{\prod_{k=1}^n (1 - L\gamma_k)\right\}^2 \sum_{i=1}^n \gamma_i (1 - L\gamma_i)^{-1}\right)^{-d/2} ,$$

where ϑ_U *is given by* (52).

Proof of Theorem 11. We bound the two terms of the right hand side of (18). The first term is dealt with the same reasonning as for the proof of Theorem 8. Regarding the second term, by [1, Theorem 1.4], π satisfies a Poincaré inequality with constant $\log^{-1}(\kappa)$. Then, the claimed bound follows from (28) and Corollary 23. □

3.8. Proof of Proposition 12

For all $x \in \mathbb{R}^d$, we have

$$\begin{aligned}\mathscr{A}^{\mathrm{L}} W_{\mathrm{c}}(x) = (\eta/4)(\|x - x^\star\|^2 + 1)^{-1/2} W_{\mathrm{c}}(x) \{(\eta/4)(\|x - x^\star\|^2 + 1)^{-1/2} \|x - x^\star\|^2 \\ - \langle \nabla U(x), x - x^\star\rangle - (\|x - x^\star\|^2 + 1)^{-1} \|x - x^\star\|^2 + d\}\end{aligned}$$

By (30), $\langle \nabla U(x), x - x^\star\rangle \geq \eta \|x - x^\star\|$ for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq M_\eta$. Then, for all x, $\|x - x^\star\| \geq K = \max(M_\eta, 4d/\eta, 1)$, $\mathscr{A}^{\mathrm{L}} W_{\mathrm{c}}(x) \leq -(\eta^2/8) W_{\mathrm{c}}(x)$, and $\sup_{\{x\in\mathcal{E}\}} \mathscr{A}^{\mathrm{L}} W_{\mathrm{c}}(x) \leq \beta$.

3.9. Proof of Proposition 13

Set $\chi = \eta/4$ and for all $x \in \mathbb{R}^d$, $\phi(x) = (\|x - x^\star\|^2 + 1)^{1/2}$. Since ϕ is 1-Lipschitz, we have by the log-Sobolev inequality [6, Theorem 5.5] for all $x \in \mathbb{R}^d$,

$$R_\gamma W_{\mathrm{c}}(x) \leq \mathrm{e}^{\chi R_\gamma \phi(x) + \chi^2\gamma} \leq \mathrm{e}^{\chi\sqrt{\|x - \gamma\nabla U(x) - x^\star\|^2 + 2\gamma d + 1} + \chi^2\gamma} . \tag{53}$$

Under **L**1 since U is convex and $x^\star$ is a minimizer of U, [29, Theorem 2.1.5 Equation (2.1.7)] shows that for all $x \in \mathbb{R}^d$,

$$\langle \nabla U(x), x - x^\star\rangle \geq (2L)^{-1} \|\nabla U(x)\|^2 + \eta \|x - x^\star\| 1_{\{\|x - x^\star\| \geq M_\eta\}} ,$$

which implies that for all $x \in \mathbb{R}^d$ and $\gamma \in \left(0, L^{-1}\right]$, we have

$$\|x - \gamma \nabla U(x) - x^\star\|^2 \leq \|x - x^\star\|^2 - 2\gamma\eta \|x - x^\star\| 1_{\{\|x-x^\star\| \geq M_\eta\}} . \tag{54}$$

Using this inequality and for all $u \in [0,1]$, $(1-u)^{1/2} - 1 \leq -u/2$, we have for all $x \in \mathbb{R}^d$, satisfiyng $\|x - x^\star\| \geq R_{\mathrm{c}} = \max(1, 2d\eta^{-1}, M_\eta)$,

$$\begin{aligned}
&\left(\|x - \gamma \nabla U(x) - x^\star\|^2 + 2\gamma d + 1\right)^{1/2} - \phi(x) \\
&\qquad \leq \phi(x) \left\{ \left(1 - 2\gamma\phi^{-2}(x)(\eta \|x - x^\star\| - d)\right)^{1/2} - 1 \right\} \\
&\qquad \leq -\gamma\phi^{-1}(x)(\eta \|x - x^\star\| - d) \leq -(\eta\gamma/2) \|x - x^\star\| \phi^{-1}(x) \leq -2^{-3/2}\eta\gamma .
\end{aligned}$$

Combining this inequality and (53), we get for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq R_{\mathrm{c}}$,

$$R_\gamma W_{\mathrm{c}}(x)/W_{\mathrm{c}}(x) \leq \mathrm{e}^{\gamma\chi(\chi - 2^{-3/2}\eta)} = \lambda .$$

By (54) and and the inequality for all $a, b \geq 0$, $\sqrt{a+1+b} - \sqrt{1+b} \leq a/2$, we get for all $x \in \mathbb{R}^d$,

$$\sqrt{\|x - \gamma \nabla U(x) - x^\star\|^2 + 2\gamma d + 1} - \phi(x) \leq \gamma d .$$

Then using this inequality in (53), we have for all $x \in \mathbb{R}^d$,

$$R_\gamma W_{\mathrm{c}}(x) \leq \lambda^\gamma W_{\mathrm{c}}(x) + \left(\mathrm{e}^{\chi\gamma(d+\chi)} - \lambda^\gamma\right) \mathrm{e}^{\eta(R_{\mathrm{c}}^2+1)^{1/2}/4} 1_{\mathrm{B}(x^\star, R_{\mathrm{c}})}(x) ,$$

Using the inequality for all $t \geq 0$, $\mathrm{e}^t - 1 \leq t\mathrm{e}^t$ concludes the proof.

3.10. Proof of Theorem 14

We preface the proof by two Lemmas.

Lemma 24. *Assume* ***L*** *1 and that U is convex. Let $(\gamma_k)_{k \in \mathbb{N}^*}$ be a nonincreasing sequence with $\gamma_1 \leq L^{-1}$. For all $n \geq 0$ and $x \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} \|y - x^\star\|^2 Q_\gamma^n(x, \mathrm{d}y) \leq \left\{4\eta^{-1}\left[1 + \log\left\{G(\lambda, c(\gamma_1), W_{\mathrm{c}}(x))\right\}\right]\right\}^2 ,$$

where $W_{\mathrm{c}}, \lambda, c$ are given in (31) and Proposition 13 respectively.

Proof. Let $n \geq 0$ and $x \in \mathbb{R}^d$. Consider the function $\phi : \mathbb{R} \to \mathbb{R}$ defined by for all $t \in \mathbb{R}$, $\phi(t) = \exp\left\{(\eta/4)(t + (4/\eta)^2)^{1/2}\right\}$. Since this function is convex on $\mathbb{R}_+$, we have by the Jensen inequality and the inequality for all $t \geq 0$, $\phi(t) \leq \mathrm{e}^{1+(\eta/4)(t+1)^{1/2}}$,

$$\phi\left(\int_{\mathbb{R}^d} \|y - x^\star\|^2 Q_\gamma^n(x, \mathrm{d}y)\right) \leq \mathrm{e}^1 Q_\gamma^n W_{\mathrm{c}}(x) .$$

The proof is then completed using Proposition 13, Lemma 1 and that ϕ is ono-to-one with for all $t \geq 1$, $\phi^{-1}(t) \leq \left(4\eta^{-1}\log(t)\right)^2$. □

Lemma 25. *Assume* ***L*** *1 and that U is convex. Then*

$$\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)} \mathrm{d}y \leq \left(\frac{(2\pi)^{(d+1)/2}(d-1)!}{\eta^d \boldsymbol{\Gamma}((d+1)/2)} + \frac{\pi^{d/2} M_\eta^d}{\boldsymbol{\Gamma}(d/2+1)} \right) . \tag{55}$$

Proof. By (30) and $U(x^\star) = 0$, we have

$$\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)} \mathrm{d}y \leq \int_{\mathbb{R}^d} \mathrm{e}^{-\eta \|y - x^\star\|} \mathrm{d}y + \int_{\mathbb{R}^d} 1_{\{\|y-x^\star\| \leq M_\eta\}} \mathrm{d}y \, .$$

Then the proof is concluded using the spherical coordinates. □

Proof of Theorem 14. By [4, Theorem 1.2], π satisfies a Poincaré inequality with constant $\log^{-1}(\kappa)$. Therefore, the second term in (18) is dealt as in the proof of Theorem 11 using (28), Lemma 25 and Lemma 22. It remains to bound $A(\gamma, x)$. Using $\nabla U(x^\star) = 0$, **L**1 and Lemma 24, we have for all $k \geq 0$,

$$\int_{\mathbb{R}^d} \|\nabla U(y)\|^2 \, Q_\gamma^k(x, \mathrm{d}y) \leq L^2 \left(4\eta^{-1} \left\{1 + \log\left\{G(\lambda, c, W_{\mathrm{c}}(x))\right\}\right\}\right)^2 \, .$$

□

3.11. Proof of Proposition 16

Under **L**1, using that $\nabla U(x^\star) = 0$, we get for all $x \in \mathbb{R}^d$,

$$\begin{aligned} \int_{\mathbb{R}^d} \|y - x^\star\|^2 \, R_\gamma(x, \mathrm{d}y) &= \|x - x^\star + \gamma(\nabla U(x^\star) - \nabla U(x))\|^2 + 2\gamma d \\ &\leq (1 + (L\gamma)^2) \, \|x - x^\star\|^2 - 2\gamma \, \langle \nabla U(x) - \nabla U(x^\star), x - x^\star \rangle + 2\gamma d \, . \end{aligned} \tag{56}$$

Then for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq M_{\mathrm{s}}$, we get using for all $t \geq 0$, $1 - t \leq \mathrm{e}^{-t}$

$$\int_{\mathbb{R}^d} \|y - x^\star\|^2 \, R_\gamma(x, \mathrm{d}y) \leq \lambda^\gamma \, \|x - x^\star\|^2 + 2\gamma d \, .$$

Using again (56) and the convexity of U, it yields for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \leq M_{\mathrm{s}}$,

$$\int_{\mathbb{R}^d} \|y - x^\star\|^2 \, R_\gamma(x, \mathrm{d}y) \leq \gamma \mathrm{c} \, ,$$

which concludes the proof.

3.12. Proof of Proposition 18

We preface the proof by a lemma.

Lemma 26. *Assume **H**4. Then, for all $x \in \mathbb{R}^d$,*

$$\|x - \gamma \nabla U(x) - x_1^\star\|^2 \leq (1 - \varpi\gamma/2) \, \|x - x_1^\star\|^2 + \gamma(\gamma + 2\varpi^{-1}) \, \|\nabla U_2\|_\infty^2 \, .$$

Proof. Using that for all $y, z \in \mathbb{R}^d$, $\|y + z\|^2 \leq (1 + \varpi\gamma/2) \, \|y\|^2 + (1 + 2(\varpi\gamma)^{-1}) \, \|z\|^2$, we get under **H**4-(b):

$$\|x - \gamma \nabla U(x) - x_1^\star\|^2 \leq (1 + \varpi\gamma/2) \, \|x - \gamma \nabla U_1(x) - x_1^\star\|^2 + \gamma(\gamma + 2\varpi^{-1}) \, \|\nabla U_2\|_\infty^2 \, .$$

Using that $\nabla U_1(x_1^\star) = 0$ and (37) in this inequality concludes the proof. □

Proof of Proposition 18. For any $\gamma \in (0, 2/(m + L_1))$, we have for all $x \in \mathbb{R}^d$:

$$\int_{\mathbb{R}^d} \|y - x_1^\star\|^2 R_\gamma(x, \mathrm{d}y) = \|x - \gamma \nabla U(x) - x_1^\star\|^2 + 2\gamma d \,. \tag{57}$$

By Lemma 26, it yields:

$$\int_{\mathbb{R}^d} \|y - x_1^\star\|^2 R_\gamma(x, \mathrm{d}y) \le (1 - \varpi\gamma/2) \|x - x_1^\star\|^2 + \gamma \left\{ (\gamma + 2\varpi^{-1}) \|\nabla U_2\|_\infty^2 + 2d \right\} \,.$$

Since $\gamma_1 \le 2/(m + L_1)$ and $(\gamma_k)_{k\ge1}$ is nonincreasing, by a straightforward induction, for $p \ge 1$ and $x \in \mathbb{R}^d$,

$$\begin{aligned}\int_{\mathbb{R}^d} \|y - x_1^\star\|^2 Q_\gamma^p(x, \mathrm{d}y) \le \prod_{k=1}^p (1 - \varpi\gamma_k/2) \|x - x_1^\star\|^2 \\ + ((\gamma_1 + 2\varpi^{-1}) \|\nabla U_2\|_\infty^2 + 2d) \sum_{i=n}^p \prod_{k=i+1}^p (1 - \varpi\gamma_k/2)\gamma_i \,,\end{aligned} \tag{58}$$

Consider the second term in the right hand side of (58). Since $\gamma_1 \le 2/(m + L_1)$, $m \le L_1$ and $(\gamma_k)_{k\ge1}$ is nonincreasing, $\max_{k\ge1} \gamma_k \le \varpi^{-1}$ and therefore:

$$\sum_{i=n}^p \prod_{k=i+1}^p (1 - \varpi\gamma_k/2)\gamma_i \le \varpi^{-1} \sum_{i=n}^p \left\{ \prod_{k=i+1}^p (1 - \varpi\gamma_k/2) - \prod_{k=i}^p (1 - \varpi\gamma_k/2) \right\} \le 2\varpi^{-1} \,.$$

□

3.13. Proof of Theorem 19

We preface the proof of the Theorem by a preliminary lemma.

Lemma 27. *Assume* ***H4****. Let $\gamma \in (0, 2/(m + L_1))$, then for all $x \in \mathbb{R}^d$,*

$$\begin{aligned}\mathrm{Ent}_\pi \left(\mathrm{d}\delta_x R_\gamma / \mathrm{d}\pi\right) \le (L_1/2) \left\{ (1 - \varpi\gamma/2) \|x - x_1^\star\|^2 + \gamma(\gamma + 2\varpi^{-1}) \|\nabla U_2\|_\infty^2 \right\} \\ + \mathrm{osc}_{\mathbb{R}^d}(U_2) - (d/2)(1 + \log(2\gamma m) - 2L_1\gamma) \,.\end{aligned}$$

Proof. Let $\gamma \in (0, 2/(m + L_1))$ and r_γ be the transition density of R_γ given by (49). Under **H** 4-(a) by [29, Theorems 2.1.8-2.1.9], we have for all $x \in \mathbb{R}^d$,

$$U_1(x) \le U_1(x_1^\star) + (L_1/2) \|x - x_1^\star\|^2 \,. \tag{59}$$

Therefore we have for all $x \in \mathbb{R}^d$

$$\mathrm{Ent}_\pi \left(\mathrm{d}\delta_x R_\gamma / \mathrm{d}\pi\right) = \int_{\mathbb{R}^d} \log(r_\gamma(x, y)/\pi(y)) r_\gamma(x, y) \mathrm{d}x \le A - (d/2)(1 + \log(4\pi\gamma)) \,, \tag{60}$$

where

$$A = \int_{\mathbb{R}^d} \left\{ U_2(y) + U_1(x_1^\star) + (L_1/2) \|y - x_1^\star\|^2 + \log\left(\int_{\mathbb{R}^d} \mathrm{e}^{-U(z)} \mathrm{d}z \right) \right\} r_\gamma(x, y) \mathrm{d}y \,.$$

By **H**4-(b) and Lemma 26, we get:

$$A \leq (L_1/2)\left\|x-\gamma\nabla U(x)-x_1^\star\right\|^2+\log\left(\int_{\mathbb{R}^d}\mathrm{e}^{-U_1(z)+U_1(x_1^\star)}\mathrm{d}z\right)+\mathrm{osc}_{\mathbb{R}^d}(U_2)+dL_1\gamma$$
$$\leq (L_1/2)\left\{(1-\varpi\gamma/2)\left\|x-x_1^\star\right\|^2+\gamma(\gamma+2\varpi^{-1})\left\|\nabla U_2\right\|_\infty^2\right\}+\mathrm{osc}_{\mathbb{R}^d}(U_2)+dL_1\gamma\,,$$

Plugging this bound in (60) gives the desired result. □

Proof of Theorem 19. We first deal with the second term in the right hand side of (18). Under **H**4, [3, Corollary 5.7.2] and the Holley-Stroock pertubation principle [16, p. 1184] show that π satisfies a log-Sobolev inequality with constant $C_{\mathrm{LS}}=-\log^{-1}(\kappa)$. So by (36) we have

$$\left\|\delta_x Q_\gamma^n P_t^{\mathrm{L}}-\pi\right\|_{\mathrm{TV}}\leq\kappa^t\left\{2\,\mathrm{Ent}_\pi\left(\frac{\mathrm{d}\delta_x Q_\gamma^n}{\mathrm{d}\pi}\right)\right\}^{1/2}.$$

We now bound $\mathrm{Ent}_\pi\left(\mathrm{d}\delta_x Q_\gamma^n/\mathrm{d}\pi\right)$ which will imply the upper bound of $C(\delta_x Q_\gamma^n)$. We proceed by induction. For $n=1$, it is Lemma 27. For $n\geq 2$, by (50) and the Jensen inequality applied to the convex function $t\mapsto t\log(t)$, we have for all $x\in\mathbb{R}^d$ and $n\geq 1$,

$$\begin{aligned}&\mathrm{Ent}_\pi\left(\mathrm{d}\delta_x Q_\gamma^n/\mathrm{d}\pi\right)\\&=\int_{\mathbb{R}^d}\log\left\{\pi^{-1}(y)\int_{\mathbb{R}^d}q_\gamma^{n-1}(x,z)r_{\gamma_n}(z,y)\mathrm{d}z\right\}\int_{\mathbb{R}^d}q_\gamma^{n-1}(x,z)r_{\gamma_n}(z,y)\mathrm{d}z\mathrm{d}y\\&\leq\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\log\left\{r_{\gamma_n}(z,y)\pi^{-1}(y)\right\}q_\gamma^{n-1}(x,z)r_{\gamma_n}(z,y)\mathrm{d}z\mathrm{d}y\,.\end{aligned}\tag{61}$$

Using Fubini's theorem, Lemma 27, Proposition 18, and the inequality $t\geq 0$, $1-t\leq\mathrm{e}^{-t}$ in (61) concludes the proof of (39).

Finally, $A(\gamma,x)$ is bounded using the inequality for all $y,z\in\mathbb{R}^d$, $\|y+z\|^2\leq 2(\|y\|^2+\|z\|^2)$, **H**4 and Proposition 18. □

4. Quantitative convergence bounds in total variation for diffusions

In this part, we derived quantitative convergence results in total variation norm for d-dimensional SDEs of the form

$$\mathrm{d}\mathbf{X}_t=b(\mathbf{X}_t)\mathrm{d}t+\mathrm{d}B_t^d\,,\tag{62}$$

started at $\mathbf{X}_0$, where $(B_t^d)_{t\geq 0}$ is a d-dimensional standard Brownian motion and $b:\mathbb{R}^d\to\mathbb{R}^d$ satisfies the following assumptions.

G1. *b is Lipschitz and for all $x,y\in\mathbb{R}^d$, $\langle b(x)-b(y),x-y\rangle\leq 0$.*

Under **G**1, [17, Theorems 2.4-3.1-6.1, Chapter IV] imply that there exists a unique solution $(\mathbf{X}_t)_{t\geq 0}$ to (62) for all initial point $x\in\mathbb{R}^d$, which is strongly Markovian. Denote by $(\mathbf{P}_t)_{t\geq 0}$ the transition semigroup associated with (62). To derive explicit bound for $\|\mathbf{P}_t(x,\cdot)-\mathbf{P}_t(y,\cdot)\|_{\mathrm{TV}}$, we use the coupling by reflection, introduced in [24] to show convergence in total variation norm for solution of SDE, and recently used by [12] to obtain exponential convergence in the Wasserstein distance of order 1. This coupling is defined as (see [8, Example 3.7]) the unique strong Markovian process $(\mathbf{X}_t,\mathbf{Y}_t)_{t\geq 0}$ on $\mathbb{R}^{2d}$, solving the SDE:

$$\begin{cases}\mathrm{d}\mathbf{X}_t &= b(\mathbf{X}_t)\mathrm{d}t+\mathrm{d}B_t^d\\ \mathrm{d}\mathbf{Y}_t &= b(\mathbf{Y}_t)\mathrm{d}t+(\mathrm{Id}-2e_te_t^T)\mathrm{d}B_t^d\,,\end{cases}\qquad\text{where } e_t=e(\mathbf{X}_t-\mathbf{Y}_t)\tag{63}$$

with $e(z) = z/\|z\|$ for $z \neq 0$ and $e(0) = 0$ otherwise. Define the coupling time

$$\tau_c = \inf\{s \geq 0 \mid \mathbf{X}_s \neq \mathbf{Y}_s\} \,. \tag{64}$$

By construction $\mathbf{X}_t = \mathbf{Y}_t$ for $t \geq \tau_c$. We denote in the sequel by $\tilde{\mathbb{P}}_{(x,y)}$ and $\tilde{\mathbb{E}}_{(x,y)}$ the probability and the expectation associated with the SDE (63) started at $(x, y) \in \mathbb{R}^{2d}$ on the canonical space of continuous function from $\mathbb{R}_+$ to $\mathbb{R}^{2d}$. We denote by $(\widetilde{\mathcal{F}}_t)_{t\geq 0}$ the canonical filtration. Since $\bar{B}_t^d = \int_0^t (\mathrm{Id} - 2e_s e_s^T)\mathrm{d}B_s^d$ is a d-dimensional Brownian motion, the marginal processes $(\mathbf{X}_t)_{t\geq 0}$ and $(\mathbf{Y}_t)_{t\geq 0}$ are under $\tilde{\mathbb{P}}_{(x,y)}$ weak solutions to (62) started at x and y respectively. The results in [24] are derived under less stringent conditions than **G** 1, but do not provide quantitative estimates.

Proposition 28 ([24, Example 5]). *Assume **G** 1 and let $(\mathbf{X}_t, \mathbf{Y}_t)_{t\geq 0}$ be the solution of (63). Then for all $t \geq 0$ and $x, y \in \mathbb{R}^d$, we have*

$$\tilde{\mathbb{P}}_{(x,y)}(\tau_c > t) = \tilde{\mathbb{P}}_{(x,y)}(\mathbf{X}_t \neq \mathbf{Y}_t) \leq 2\left(\mathbf{\Phi}\left\{\left(2t^{1/2}\right)^{-1}\|x - y\|\right\} - 1/2\right) \,.$$

Proof. For $t < \tau_c$, $\mathbf{X}_t - \mathbf{Y}_t$ is the solution of the SDE

$$\mathrm{d}\{\mathbf{X}_t - \mathbf{Y}_t\} = \{b(\mathbf{X}_t) - b(\mathbf{Y}_t)\}\,\mathrm{d}t + 2e_t \mathrm{d}B_t^1 \,,$$

where $B_t^1 = \int_0^t 1_{\{s<\tau_c\}} e_s^T \mathrm{d}B_s^d$. Using the Itô's formula and **G**1, we have for all $t < \tau_c$,

$$\|\mathbf{X}_t - \mathbf{Y}_t\| = \|x - y\| + \int_0^t \langle b(\mathbf{X}_s) - b(\mathbf{Y}_s), e_s\rangle\,\mathrm{d}s + 2B_t^1 \leq \|x - y\| + 2B_t^1 \,.$$

Therefore, for all $x, y \in \mathbb{R}^d$ and $t \geq 0$, we get

$$\begin{aligned}\tilde{\mathbb{P}}_{(x,y)}(\tau_c > t) &\leq \tilde{\mathbb{P}}_{(x,y)}\left(\min_{0\leq s\leq t} B_s^1 \geq \|x - y\|/2\right) \\ &= \tilde{\mathbb{P}}_{(x,y)}\left(\max_{0\leq s\leq t} B_s^1 \leq \|x - y\|/2\right) = \tilde{\mathbb{P}}_{(x,y)}\left(|B_t^1| \leq \|x - y\|/2\right) \,.\end{aligned}$$

where we have used the reflection principle in the last identity. □

Define for $R > 0$ the set $\Delta_R = \{x, y \in \mathbb{R}^d \mid \|x - y\| \leq R\}$. Proposition 28 and Lindvall's inequality give that, for all $\epsilon \in (0, 1)$ and $t \geq \omega(\epsilon, R)$,

$$\sup_{(x,y)\in\Delta_R} \|\mathbf{P}_t(x, \cdot) - \mathbf{P}_t(y, \cdot)\|_{\mathrm{TV}} \leq (1 - \epsilon) \,, \tag{65}$$

where ω is defined in (34). To obtain quantitative exponential bounds in total variation for any $x, y \in \mathbb{R}^d$, it is required to control some exponential moments of the successive return times to Δ_R. This is first achieved by using a drift condition for the generator $\mathscr{A}$ associated with the SDE (62) defined for all $f \in C^2(\mathbb{R}^d)$ by

$$\mathscr{A}f = \langle b, \nabla f\rangle + (1/2)\Delta f \,.$$

Consider the following assumption:

G 2. *(i) There exist a twice continuously differentiable function $V : \mathbb{R}^d \mapsto [1, \infty)$ and constants $\theta > 0$, $\beta \geq 0$ such that*

$$\mathscr{A}V \leq -\theta V + \beta \,. \tag{66}$$

(ii) There exists $\delta > 0$ and $\mathrm{R} > 0$ such that $\Theta \subset \Delta_{\mathrm{R}}$ where

$$\Theta = \{(x, y) \in \mathbb{R}^{2d} \mid V(x) + V(y) \leq 2\theta^{-1}\beta + \delta\} \,. \tag{67}$$

For $t > 0$, and G a closed subset of $\mathbb{R}^{2d}$, define by $\mathrm{T}_1^{G,t}$ the first return time to G delayed by t:

$$\mathrm{T}_1^{G,t} = \inf\{s \geq t \mid (\mathbf{X}_s, \mathbf{Y}_s) \in G\} \,.$$

For $j \geq 2$, define recursively the j-th return times to G delayed by t by

$$\mathrm{T}_j^{G,t} = \inf\{s \geq \mathrm{T}_{j-1}^{G,t} + t \mid (\mathbf{X}_s, \mathbf{Y}_s) \in G\} = \mathrm{T}_{j-1}^{G,t} + \mathrm{T}_1^{G,t} \circ \mathrm{S}_{\mathrm{T}_{j-1}^{G,t}} \,, \tag{68}$$

where S is the shift operator on the canonical space. By [13, Proposition 1.5 Chapter 2], the sequence $(\mathrm{T}_j^{G,t})_{j\geq 1}$ is a sequence of stopping time with respect to $(\widetilde{\mathcal{F}}_t)_{t\geq 0}$.

Proposition 29. *Assume **G**1 and **G**2. For all $x, y \in \mathbb{R}^d$, $\epsilon \in (0, 1)$ and $j \geq 1$, we have*

$$\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{\tilde{\theta}\mathrm{T}_j^{\Theta,\omega(\epsilon,\mathrm{R})}}\right] \leq \{\mathrm{K}(\epsilon)\}^{j-1}\left\{(1/2)(V(x) + V(y)) + \mathrm{e}^{\tilde{\theta}\omega(\epsilon,\mathrm{R})}\tilde{\theta}^{-1}\beta\right\} \,,$$

$$\tilde{\theta} = \theta^2\delta(2\beta + \theta\delta)^{-1} \,,\ \mathrm{K}(\epsilon) = \tilde{\theta}^{-1}\beta\left(1 + \mathrm{e}^{\tilde{\theta}\omega(\epsilon,\mathrm{R})}\right) + \delta/2 \,, \tag{69}$$

where ω is defined in (34).

Proof. For notational simplicity, set $\mathrm{T}_j = \mathrm{T}_j^{\Theta,\omega(\epsilon,\mathrm{R})}$. Note that for all $x, y \in \mathbb{R}^d$,

$$\mathscr{A}V(x) + \mathscr{A}V(y) \leq -\tilde{\theta}(V(x) + V(y)) + 2\beta 1_\Theta(x, y)$$

Then by the Dynkin formula (see e.g. [28, Eq. (8)]) the process

$$t \mapsto (1/2)\mathrm{e}^{\tilde{\theta}(\mathrm{T}_1\wedge t)}\left\{V\left(\mathbf{X}_{\mathrm{T}_1\wedge t}\right) + V\left(\mathbf{Y}_{\mathrm{T}_1\wedge t}\right)\right\} \,, \quad t \geq \omega(\epsilon, \mathrm{R}) \,,$$

is a positive supermartingale. Using the optional stopping theorem and the Markov property, we have, using that for all $t \geq 0$ $\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{\tilde{\theta}t}V(X_t)\right] \leq V(x) + \beta\tilde{\theta}^{-1}\mathrm{e}^{\tilde{\theta}t}$,

$$\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{\tilde{\theta}\mathrm{T}_1}\right] \leq (1/2)(V(x) + V(y)) + \mathrm{e}^{\tilde{\theta}\omega(\epsilon,\mathrm{R})}\tilde{\theta}^{-1}\beta \,.$$

The result then follows from this inequality and the strong Markov property. □

Theorem 30. *Assume **G**1 and **G**2. Then for all $\epsilon \in (0, 1)$, $t \geq 0$ and $x, y \in \mathbb{R}^d$,*

$$\|\mathbf{P}_t(x, \cdot) - \mathbf{P}_t(y, \cdot)\|_{\mathrm{TV}} \leq ((1 - \epsilon)^{-2} + (1/2)\{V(x) + V(y)\})\kappa^t \,,$$

where ω are defined in (34), $\tilde{\theta}, \mathrm{K}(\epsilon)$ in (69) and

$$\log(\kappa) = \tilde{\theta}\log(1 - \epsilon)\{\log(\mathrm{K}(\epsilon)) - \log(1 - \epsilon)\}^{-1} \,.$$

Proof. Let $x, y \in \mathbb{R}^d$ and $t \geq 0$. For all $\ell \geq 1$ and $\epsilon \in (0, 1)$,

$$\tilde{\mathbb{P}}_{(x,y)}(\tau_c > t) \leq \tilde{\mathbb{P}}_{(x,y)}(\tau_c > t, \mathrm{T}_\ell \leq t) + \tilde{\mathbb{P}}_{(x,y)}(\mathrm{T}_\ell > t) \,, \tag{70}$$

where $\mathrm{T}_\ell = \mathrm{T}_\ell^{\Theta,\omega(\epsilon,\mathrm{R})}$. We now bound the two terms in the right hand side of this equation. For the first term, since $\Theta \subset \Delta_{\mathrm{R}}$, by (65), we have conditioning successively on $\widetilde{\mathcal{F}}_{\mathrm{T}_j}$, for $j = \ell, \ldots, 1$, and using the strong Markov property,

$$\tilde{\mathbb{P}}_{(x,y)}(\tau_c > t, \mathrm{T}_\ell \leq t) \leq (1 - \epsilon)^\ell \,. \tag{71}$$

For the second term, using Proposition 29 and the Markov inequality, we get

$$\tilde{\mathbb{P}}_{(x,y)}\left(\mathrm{T}_\ell > t\right) \leq \mathrm{e}^{-\tilde{\theta}t}\{\mathrm{K}(\epsilon)\}^{\ell-1}\left\{(1/2)(V(x)+V(y)) + \mathrm{e}^{\tilde{\theta}\omega(\epsilon,\mathrm{R})}\tilde{\theta}^{-1}\beta\right\} \tag{72}$$

Combining (71)-(72) in (70) and taking $\ell = \left\lfloor \tilde{\theta}t \big/ (\log(\mathrm{K}(\epsilon)) - \log(1-\epsilon)) \right\rfloor$ concludes the proof. □

More precise bounds can be obtained under more stringent assumption on the drift b; see [5] and [12].

G3. *There exist $\tilde{M}_\mathrm{s} \geq 1$ and $\tilde{m}_\mathrm{s} > 0$, such that for all $x, y \in \mathbb{R}^d$, $\|x-y\| \geq \tilde{M}_\mathrm{s}$,*

$$\langle b(x) - b(y), x-y\rangle \leq -\tilde{m}_\mathrm{s}\|x-y\|^2 \ .$$

Proposition 31. *Assume **G**1 and **G**3.*

(a) For all $x, y \in \mathbb{R}^d$ and $\epsilon \in (0,1)$

$$\tilde{\mathbb{E}}_{(x,y)}\left[\exp\left((\tilde{m}_\mathrm{s}/2)\left(\tau_c \wedge \mathrm{T}_1^{\Delta_{\tilde{M}_\mathrm{s}},\omega(\epsilon,\tilde{M}_\mathrm{s})}\right)\right)\right] \leq 1 + \|x-y\| + (1+\tilde{M}_\mathrm{s})\mathrm{e}^{\tilde{m}_\mathrm{s}\omega(\epsilon,\tilde{M}_\mathrm{s})/2} \ .$$

(b) For all $x, y \in \mathbb{R}^d$, $\epsilon \in (0,1)$ and $j \geq 1$

$$\tilde{\mathbb{E}}_{(x,y)}\left[\exp\left((\tilde{m}_\mathrm{s}/2)\left(\tau_c \wedge \mathrm{T}_j^{\Delta_{\tilde{M}_\mathrm{s}},\omega(\epsilon,\tilde{M}_\mathrm{s})}\right)\right)\right]$$
$$\leq \{\mathrm{D}(\epsilon)\}^{j-1}\left\{1 + \|x-y\| + (1+\tilde{M}_\mathrm{s})\mathrm{e}^{\tilde{m}_\mathrm{s}\omega(\epsilon,\tilde{M}_\mathrm{s})/2}\right\} \ ,$$

$$\mathrm{D}(\epsilon) = (1+\mathrm{e}^{\tilde{m}_\mathrm{s}\omega(\epsilon,\tilde{M}_\mathrm{s})/2})(1+\tilde{M}_\mathrm{s}) \ , \tag{73}$$

where ω is given in (34).

Proof. In the proof, we set $\mathrm{T}_j = \mathrm{T}_j^{\Delta_{\tilde{M}_\mathrm{s}},\omega(\epsilon,\tilde{M}_\mathrm{s})}$.

(a) Consider the sequence of increasing stopping time

$$\tau_k = \inf\{t > 0 \mid \|\mathbf{X}_t - \mathbf{Y}_t\| \notin \left[k^{-1}, k\right]\} \ , \quad k \geq 1 \ ,$$

and set $\zeta_k = \tau_k \wedge \mathrm{T}_1$. We derive a bound on $\tilde{\mathbb{E}}_{(x,y)}[\exp\{(\tilde{m}_\mathrm{s}/2)\zeta_k\}]$ independent on k. Since $\lim_{k\to+\infty} \uparrow \tau_k = \tau_c$ almost surely, the monotone convergence theorem implies that the same bound holds for $\tilde{\mathbb{E}}_{(x,y)}[\exp\{(\tilde{m}_\mathrm{s}/2)(\tau_c \wedge \mathrm{T}_1)\}]$. Set now $\mathrm{W}_\mathrm{s}(x,y) = 1 + \|x-y\|$. Since $\mathrm{W}_\mathrm{s} \geq 1$ and $\tau_c < \infty$ a.s by Proposition 28, it suffices to give a bound on $\tilde{\mathbb{E}}_{(x,y)}[\exp\{(\tilde{m}_\mathrm{s}/2)\zeta_k\}\mathrm{W}_\mathrm{s}(\mathbf{X}_{\zeta_k}, \mathbf{Y}_{\zeta_k})]$. By Itô's formula, we have for all $v, t \leq \tau_c$, $v \leq t$

$$\mathrm{e}^{\tilde{m}_\mathrm{s}t/2}\mathrm{W}_\mathrm{s}(\mathbf{X}_t, \mathbf{Y}_t) = \mathrm{e}^{\tilde{m}_\mathrm{s}v/2}\mathrm{W}_\mathrm{s}(\mathbf{X}_v, \mathbf{Y}_v) + (\tilde{m}_\mathrm{s}/2)\int_v^t \mathrm{e}^{\tilde{m}_\mathrm{s}u/2}\mathrm{W}_\mathrm{s}(\mathbf{X}_u, \mathbf{Y}_u)\mathrm{d}u$$
$$+ \int_v^t \mathrm{e}^{\tilde{m}_\mathrm{s}u/2}\langle b(\mathbf{X}_u) - b(\mathbf{Y}_u), e_u\rangle \,\mathrm{d}u + 2\int_v^t \mathrm{e}^{\tilde{m}_\mathrm{s}u/2}\mathrm{d}B_u^1 \ . \tag{74}$$

Using **G**3(*b*), we have for all $k \geq 1$ and $t_\mathrm{s} = \omega(\epsilon, \tilde{M}_\mathrm{s}) \leq v \leq t$

$$\mathrm{e}^{(\tilde{m}_\mathrm{s}/2)(\zeta_k\wedge t)}\mathrm{W}_\mathrm{s}(\mathbf{X}_{\zeta_k\wedge t}, \mathbf{Y}_{\zeta_k\wedge t}) \leq \mathrm{e}^{(\tilde{m}_\mathrm{s}/2)(\zeta_k\wedge v)}\mathrm{W}_\mathrm{s}(\mathbf{X}_{\zeta_k\wedge v}, \mathbf{Y}_{\zeta_k\wedge v}) + 2\int_{\zeta_k\wedge v}^{\zeta_k\wedge t} \mathrm{e}^{\tilde{m}_\mathrm{s}u/2}\mathrm{d}B_u^1 \ .$$

So the process

$$\{\exp\left((\tilde{m}_{\mathrm{s}}/2)(\zeta_k \wedge t)\right) \mathrm{W}_{\mathrm{s}}(\mathbf{X}_{\zeta_k \wedge t}, \mathbf{Y}_{\zeta_k \wedge t})\}_{t \geq t_{\mathrm{s}}} \ ,$$

is a positive supermartingale and by the optional stopping theorem, we get

$$\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{(\tilde{m}_{\mathrm{s}}/2)\zeta_k}\mathrm{W}_{\mathrm{s}}(\mathbf{X}_{\zeta_k}, \mathbf{Y}_{\zeta_k})\right] \leq \tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{(\tilde{m}_{\mathrm{s}}/2)(\tau_k \wedge t_{\mathrm{s}})}\mathrm{W}_{\mathrm{s}}(\mathbf{X}_{\tau_k \wedge t_{\mathrm{s}}}, \mathbf{Y}_{\tau_k \wedge t_{\mathrm{s}}})\right] \ , \tag{75}$$

where we used that $\zeta_k \wedge t_{\mathrm{s}} = \tau_k \wedge t_{\mathrm{s}}$. By (74), **G**1 and **G**3, we have

$$\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{(\tilde{m}_{\mathrm{s}}/2)(\tau_k \wedge t_{\mathrm{s}})}\mathrm{W}_{\mathrm{s}}(\mathbf{X}_{\tau_k \wedge t_{\mathrm{s}}}, \mathbf{Y}_{\tau_k \wedge t_{\mathrm{s}}})\right] \leq \mathrm{W}_{\mathrm{s}}(x,y) + (1 + \tilde{M}_{\mathrm{s}})\mathrm{e}^{\tilde{m}_{\mathrm{s}} t_{\mathrm{s}}/2} \ ,$$

and (75) becomes

$$\tilde{\mathbb{E}}_{(x,y)}\left[\mathrm{e}^{(\tilde{m}_{\mathrm{s}}/2)\zeta_k}\mathrm{W}_{\mathrm{s}}(\mathbf{X}_{\zeta_k}, \mathbf{Y}_{\zeta_k})\right] \leq \mathrm{W}_{\mathrm{s}}(x,y) + (1 + \tilde{M}_{\mathrm{s}})\mathrm{e}^{\tilde{m}_{\mathrm{s}} t_{\mathrm{s}}/2} \ .$$

(b) The proof is by induction. The case $j = 1$ has been established above. Now let $j \geq 2$. Since on the event $\{\tau_c > \mathrm{T}_{j-1}\}$, we have

$$\tau_c \wedge \mathrm{T}_j = \mathrm{T}_{j-1} + (\tau_c \wedge \mathrm{T}_1) \circ \mathrm{S}_{\mathrm{T}_{j-1}} \ ,$$

where S is the shift operator, we have conditioning on $\widetilde{\mathcal{F}}_{\mathrm{T}_{j-1}}$, using the strong Markov property, Proposition 28 and the first part,

$$\tilde{\mathbb{E}}_{(x,y)}\left[1_{\tau_c > \mathrm{T}_{j-1}} \exp\left((\tilde{m}_{\mathrm{s}}/2)\left(\tau_c \wedge \mathrm{T}_j\right)\right)\right] \leq \mathrm{D}(\epsilon)\, \tilde{\mathbb{E}}_{(x,y)}\left[1_{\tau_c > \mathrm{T}_{j-1}} \exp\left((\tilde{m}_{\mathrm{s}}/2)\mathrm{T}_{j-1}\right)\right] \ ,$$

Then the proof follows since $\mathrm{D}(\epsilon) \geq 1$.

□

Theorem 32. *Assume **G**1 and **G**3. Then for all $\epsilon \in (0,1)$, $t \geq 0$ and $x, y \in \mathbb{R}^d$,*

$$\|\mathbf{P}_t(x,\cdot) - \mathbf{P}_t(y,\cdot)\|_{\mathrm{TV}} \leq \left\{(1-\epsilon)^{-2} + 1 + \|x - y\|\right\} \kappa^t \ ,$$

$$\log(\kappa) = (\tilde{m}_{\mathrm{s}}/2)\log(1-\epsilon)(\log(\mathrm{D}(\epsilon)) - \log(1-\epsilon))^{-1} \ .$$

where $\mathrm{D}(\epsilon)$ is defined in (73).

Proof. The proof is along the same lines as Theorem 30. Set $\mathrm{T}_j = \mathrm{T}_j^{\Delta_{\tilde{M}_{\mathrm{s}}}, \omega(\epsilon, \tilde{M}_{\mathrm{s}})}$ for $j \geq 1$. Let $x, y \in \mathbb{R}^d$ and $t \geq 0$. For all $\ell \geq 1$ and $\epsilon \in (0,1)$,

$$\tilde{\mathbb{P}}_{(x,y)}\left(\tau_c > t\right) \leq \tilde{\mathbb{P}}_{(x,y)}\left(\tau_c > t, \mathrm{T}_\ell \leq t\right) + \tilde{\mathbb{P}}_{(x,y)}\left(\mathrm{T}_\ell \wedge \tau_c > t\right) \ . \tag{76}$$

For the first term, by (65) we have conditioning successively on $\widetilde{\mathcal{F}}_{\mathrm{T}_j}$, for $j = \ell, \cdots, 1$, and using the strong Markov property,

$$\tilde{\mathbb{P}}_{(x,y)}\left(\tau_c > t, \mathrm{T}_\ell \leq t\right) \leq (1-\epsilon)^\ell \ . \tag{77}$$

For the second term, using Proposition 31-(b) and the Markov inequality, we get

$$\tilde{\mathbb{P}}_{(x,y)}\left(\mathrm{T}_\ell \wedge \tau_c > t\right) \leq \mathrm{e}^{-(\tilde{m}_{\mathrm{s}} t/2)}\{\mathrm{D}(\epsilon)\}^{\ell-1}\left\{1 + \|x-y\| + (1+\tilde{M}_{\mathrm{s}})\mathrm{e}^{\tilde{m}_{\mathrm{s}}\omega(\epsilon,\tilde{M}_{\mathrm{s}})/2}\right\} \ . \tag{78}$$

Combining (77)-(78) in (76) and taking $\ell = \left\lfloor (\tilde{m}_{\mathrm{s}} t/2)/(\log(\mathrm{D}(\epsilon)) - \log(1-\epsilon))\right\rfloor$ concludes the proof. □

4.1. Proof of Theorem 15 and Theorem 17

Recall that $(P_t^{\mathrm{L}})_{t\geq 0}$ is the Markov semigroup of the Langevin equation associated with U and let $\mathscr{A}^{\mathrm{L}}$ be the corresponding generator. Since $(P_t^{\mathrm{L}})_{t\geq 0}$ is reversible with respect to π, we deduce from Theorem 30 and Theorem 32 quantitative bounds for the exponential convergence of $(P_t^{\mathrm{L}})_{t\geq 0}$ to π in total variation noting that if $(Y_t^{\mathrm{L}})_{t\geq 0}$ is a solution of (1), then $(Y_{t/2}^{\mathrm{L}})_{t\geq 0}$ is a weak solution of the rescaled Langevin diffusion:

$$\mathrm{d}\tilde{Y}_t^{\mathrm{L}} = -(1/2)\nabla U(\tilde{Y}_t^{\mathrm{L}})\mathrm{d}t + \mathrm{d}B_t^d \ . \tag{79}$$

Proof of Theorem 15. Since the generator associated with the SDE (79) is $(1/2)\mathscr{A}^{\mathrm{L}}$, Proposition 12 shows that (66) holds for W_{c} with constants $\theta/2$ and $\beta/2$. Using that for all $a_1, a_2 \in \mathbb{R}$, $\mathrm{e}^{(a_1+a_2)/2} \leq (1/2)(\mathrm{e}^{a_1} + \mathrm{e}^{a_2})$, **G**2-(ii) holds for $\delta = 2\theta^{-1}\beta$ and $\mathrm{R} = (8/\eta)\log(4\theta^{-1}\beta)$. By Theorem 30 with $\epsilon = 1/2$, we get for all $x, y \in \mathbb{R}^d$ and $t \geq 0$

$$\|P_t^{\mathrm{L}}(x,\cdot) - P_t^{\mathrm{L}}(y,\cdot)\|_{\mathrm{TV}} \leq \kappa^t \left\{4 + (1/2)(W_{\mathrm{c}}(x) + W_{\mathrm{c}}(y))\right\} \ , \tag{80}$$

where κ is defined in (35a). By [28, Theorem 4.3-(ii)], (66) implies that $\int_{\mathbb{R}^d} W_{\mathrm{c}}(y)\pi(\mathrm{d}y) \leq \beta\theta^{-1}$. The proof is then concluded using this bound, (80), that π is invariant for $(P_t^{\mathrm{L}})_{t\geq 0}$, we get for all $x \in \mathbb{R}^d$,

$$\|\delta_x P_{\Gamma_{n+1,p}}^{\mathrm{L}} - \mu\|_{\mathrm{TV}} \leq \kappa^{\Gamma_{n+1,p}} \left\{4 + (1/2)(\beta\theta^{-1} + W_{\mathrm{c}}(x))\right\} \ .$$

The proof is then concluded by Proposition 13 and Lemma 1. □

Proof of Theorem 17. By applying Theorem 32 with $\epsilon = 1/2$, the triangle inequality and using that π is invariant for $(P_t^{\mathrm{L}})_{t\geq 0}$, we have

$$\|P_t^{\mathrm{L}}(x,\cdot) - \pi\|_{\mathrm{TV}} \leq \left\{5 + \|x - x^\star\| + \int_{\mathbb{R}^d} \|y - x^\star\| \, \mathrm{d}\pi(y)\right\} \kappa^t \ .$$

It remains to show that $\int_{\mathbb{R}^d} \|y - x^\star\| \, \mathrm{d}\pi(y) \leq (d/m + M_{\mathrm{s}}^2)^{1/2}$. For this, we establish a drift inequality for the generator $\mathscr{A}^{\mathrm{L}}$ of the Langevin SDE associated with U. Consider the function $W_{\mathrm{s}}(x) = \|x - x^\star\|^2$. For all $x \in \mathbb{R}^d$, we have using $\nabla U(x^\star) = 0$,

$$\mathscr{A}^{\mathrm{L}} W_{\mathrm{s}}(x) \leq 2(d - \langle \nabla U(x) - \nabla U(x^\star), x - x^\star \rangle) \ .$$

Therefore by **G**3, for all $x \in \mathbb{R}^d$, $\|x - x^\star\| \geq M_{\mathrm{s}}$, we get

$$\mathscr{A}^{\mathrm{L}} W_{\mathrm{s}}(x) \leq -2m W_{\mathrm{s}}(x) + 2d \ ,$$

and for all $x \in \mathbb{R}^d$,

$$\mathscr{A}^{\mathrm{L}} W_{\mathrm{s}}(x) \leq -2m W_{\mathrm{s}}(x) + 2(d + m M_{\mathrm{s}}^2) \ .$$

By [28, Theorem 4.3-(ii)], we get $\int_{\mathbb{R}^d} W_{\mathrm{s}}(y)\mathrm{d}\pi(y) \leq d/m + M_{\mathrm{s}}^2$. The bound on $C(\delta_x Q_\gamma^n)$ is a consequence of the Cauchy-Schwarz inequality, Proposition 16 and Lemma 1. The bound for $A(\gamma, x)$ similarly follows from **L**1, Proposition 16 and Lemma 1

□

Acknowledgements

The authors are indebted to Arnaud Guillin for sharing his knowledge of Poincaré and log-Sobolev inequalities

References

[1] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures. *Electronic Communications in Probability [electronic only]*, 13:60–66, 2008.

[2] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.

[3] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[4] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 10 1999.

[5] François Bolley, Ivan Gentil, and Arnaud Guillin. Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations. *J. Funct. Anal.*, 263(8):2430–2457, 2012.

[6] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[7] P. Cattiaux and A. Guillin. Trends to equilibrium in total variation distance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(1):117–145, 2009.

[8] M. F. Chen and S. F. Li. Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1):151–177, 1989.

[9] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC methods for functions: modifying old algorithms to make them faster. *Statist. Sci.*, 28(3):424–446, 2013.

[10] A. Dalalyan. Theoretical guarantees for approximate sampling from a smooth and log-concave density. submitted 1412.7392, arXiv, December 2014.

[11] A. S. Dalalyan and A. B. Tsybakov. Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. System Sci.*, 78(5):1423–1443, 2012.

[12] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, pages 1–36, 2015.

[13] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, 1986.

[14] U. Grenander. Tutorial in pattern theory. Division of Applied Mathematics, Brown University, Providence.

[15] U. Grenander and M. I. Miller. Representations of knowledge in complex systems. *J. Roy. Statist. Soc. Ser. B*, 56(4):549–603, 1994. With discussion and a reply by the authors.

[16] Richard Holley and Daniel Stroock. Logarithmic sobolev inequalities and stochastic ising models. *Journal of statistical physics*, 46(5):1159–1194, 1987.

[17] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library. Elsevier Science, 1989.

[18] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.

[19] S. Kullback. *Information theory and statistics*. Dover Publications, Inc., Mineola, NY, 1997. Reprint of the second (1968) edition.

[20] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3):367–405, 2002.

[21] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion: the case of a weakly mean reverting drift. *Stoch. Dyn.*, 3(4):435–451, 2003.

[22] V. Lemaire. *Estimation de la mesure invariante d'un processus de diffusion*. PhD thesis,

Université Paris-Est, 2005.
[23] V. Lemaire and S. Menozzi. On some non asymptotic bounds for the Euler scheme. *Electron. J. Probab.*, 15:no. 53, 1645–1681, 2010.
[24] T. Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.
[25] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.
[26] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009.
[27] S. P. Meyn and R. L. Tweedie. Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Advances in Applied Probability*, 25(3), 1993.
[28] Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.
[29] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Applied Optimization. Springer, 2004.
[30] G. Parisi. Correlation functions and computer simulations. *Nuclear Physics B*, 180:378–384, 1981.
[31] G. O. Roberts and R. L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.
[32] G. O. Roberts and R. L. Tweedie. Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.*, 37(2):359–373, 2000.
[33] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.